BALLISTIC TRANSPORT AND TUNNELING IN SMALL SYSTEMS

A THESIS

Submitted to the Department of Physics
and the Institute of Engineering and Science
of Bilkent University
in Partial Fulfillment of the Requirements
for the Degree of
Doctor of Philosophy

By
A. ERKAN TEKMAN
OCTOBER, 1990

BALLISTIC TRANSPORT AND TUNNELING IN SMALL SYSTEMS

A THESIS
SUBMITTED TO THE DEPARTMENT OF PHYSICS
AND THE INSTITUTE OF ENGINEERING AND SCIENCE
OF BILKENT UNIVERSITY
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
DOCTOR OF PHILOSPHY

By
A. Erkan Tekman
October 1990

I certify that I have read this thesis and that in my opinion it is fully adequate, in scope and in quality, as a dissertation for the degree of Doctor of Philosophy.

Prof. Salim Çıracı (Supervisor)

I certify that I have read this thesis and that in my opinion it is fully adequate, in scope and in quality, as a dissertation for the degree of Doctor of Philosophy.

Prof. M. Cemal Yalabık

I certify that I have read this thesis and that in my opinion it is fully adequate, in scope and in quality, as a dissertation for the degree of Doctor of Philosophy.

Prof. Şinasi Ellialtıoğlu

I certify that I have read this thesis and that in my opinion it is fully adequate, in scope and in quality, as a dissertation for the degree of Doctor of Philosophy.

Assoc. Prof. Atilla Erçelebi

I certify that I have read this thesis and that in my opinion it is fully adequate, in scope and in quality, as a dissertation for the degree of Doctor of Philosophy.

Prof. Abdullah Atalar

Approved for the Institute of Engineering and Science:

Prof. Mehmet Baray,
Director of Institute of Engineering and Science

Abstract

BALLISTIC TRANSPORT AND TUNNELING IN SMALL SYSTEMS

A. Erkan Tekman
Ph. D. in Physics
Supervisor: Prof. Salim Çıracı
October 1990

Ballistic transport and tunneling of electrons in mesoscopic systems have become one of the most important subjects of condensed matter physics. The quantum point contacts and scanning tunneling microscope form the basic experimental tools in this area and have been used for understanding many features of small systems. In this work ballistic transport and tunneling in small systems are investigated theoretically.

Ballistic transport through narrow constrictions is investigated for a variety of configurations. It is found that for a uniform constriction the conductance is quantized in units of the quantum of conductance ($2e^2/h$) for long channels. The interference of waves in the constriction gives rise to the resonance structure superimposed on the quantized steps. The lack of the resonance structure in the experimental results are attributed to temperature effects and/or adiabatic transport due to tapering of the constriction. It is shown that elastic scattering by an impurity distorts the quantization of conductance. Novel resonant tunneling effects due to formation of bound states are predicted for an attractive impurity or a local widening at the center of the constriction.

It is shown that the probing in scanning tunneling microscopy have very much in common with narrow constrictions. The transition from tunneling to point contact regime is explained by the vanishing effective potential barrier as a result of tip-sample interaction. For noble and simple metals it is conjectured that lateral position dependent interaction between the tip and sample leads to corrugation of the potential barrier and in turn to atomic corrugation observed by scanning tunneling microscopy. The focused field emission of electrons from point sources is analyzed in a systematical way. The effective barrier due to the lateral confinement and nonadiabatic transport through the horn-like opening are found to be responsible for focusing.

The nonequilibrium nature of transport is investigated by use of Keldysh Green's function technique. The effects of elastic and inelastic scattering are analyzed in a strictly one-dimensional geometry. The features of voltage and current probes are studied and the Landauer formulae are examined for multiprobe measurements.

Keywords: Mesoscopics, quantum point contacts, ballistic transport, tunneling, quantized conductance, adiabatic transport, resonant tunneling, scanning tunneling microscopy, focused field emission, nonequilibrium quantum transport, Keldysh technique, Landauer formulae.

Özet

KÜÇÜK SİSTEMLERDE BALİSTİK İLETİM VE TÜNELLEME

A. Erkan Tekman
Fizik Doktora
Tez Yöneticisi: Prof. Dr. Salim Çıracı
Ekim 1990

Mezoskopik sistemlerde balistik iletim ve tünelleme yoğun madde fiziğinin önemli konularından biri haline gelmiştir. Kuantum nokta bağlantıları ve tarama tünelleme mikroskobu bu çalışmaların temel deney araçlarıdır ve küçük sistemlerin pekçok özelliğinin anlaşılmasında kullanılmışlardır. Bu çalışmada küçük sistemlerde balistik iletim ve tünelleme kuramsal olarak ele alındı.

Dar bağlantılarda balistik iletim çeşitli durumlar için incelendi. Düzgün bağlantılarda iletkenliğin kuantalaştığı ve uzun kanallar için $(2e^2/h)$in katları olarak değiştiği bulundu. Dalgaların girişimi sonucu kuantalaşmış basamaklar üzerinde bir rezonans yapısı gözlendi. Rezonans yapısının deneysel sonuçlarda görülmemesi sıcaklık ve/veya bağlantının yuvarlatılması sonucu ortaya çıkan adiyabatik iletim ile açıklandı. Kanal içindeki bozukluklardan elastik saçılma sonucu kuantalaşmış basamakların bozulduğu bulundu. Bağlantının merkezinde çekici bozukluklar ya da genişleme olmasının bağlı durumların oluşmasına ve rezonant tünellemeye yol açacağı öngörüldü.

Tarama tünelleme mikroskobundaki ölçümün dar bağlantılarla pekçok ortak noktası olduğu gösterildi. Tünellemeden nokta bağlantıya geçişe uç-yüzey

etkileşimi sonucu etkin potansiyel eşiğinin yok olmasının yol açtığı bulundu. Soy ve basit metallerde pozisyona bağlı uç-yüzey etkileşiminin potansiyel eşiğinin modülasyonuna ve atomsal yükselti farklarının gözlenmesine yol açtığı öngörüldü. Nokta kaynaklardan elektronların odaklanmış alan saçınımı sistematik olarak incelendi. Paralel sıkıştırma sonucu oluşan etkin potansiyelin ve boru benzeri açıklıktaki adiyabatik olmayan iletimin odaklanmaya yol açtığı bulundu.

Dengedışı iletim özellikleri Keldysh'in Green fonksiyonu tekniği ile araştırıldı. Elastik ve inelastik saçılımın etkileri tek boyutlu bir geometride çalışıldı. Gerilim ve akım ölçüm bağlantılarının önemi ve çok bağlantılı ölçümler için Landauer formülleri incelendi.

Anahtar sözcükler: Mezoskopik fizik, kuantum nokta bağlantıları, balistik iletim, tünelleme, kuantalaşmış iletkenlik, adiyabatik iletim, resonant tünelleme, tarama tünelleme mikroskobu, odaklanmış alan saçınımı, dengedışı kuantum iletimi, Keldysh tekniği, Landauer formülleri.

Acknowledgement

It is my pleasure to express my deepest gratitude to Prof. Salim Çıracı for his supervision to my graduate study. I have enjoyed working with him for the last seven years during which he has been my teacher, supervisor and collaborator. I wish to thank him for his stimulation in an enthusiastic, decisive, and yet a friendly way, his encouragement in times of despair, and his invaluable contributions in this and related works.

I wish to express my thanks to all colleagues who sent their work prior to publication and with whom I had many fruitful discussions. Among these I wish to acknowledge discussions and correspondence with Dr. Markus Büttiker, Dr. Pierre Guéret and Phil Bagwell.

I acknowledge partial support by the Joint Project Agreement between IBM Zurich Research Laboratory and Department of Physics, Bilkent University. I want to extend my thanks to colleagues at IBM Zurich for their hospitality and discussions during my visit to the laboratory. Special thanks are due to Prof. Alexis Baratoff and Dr. Inder P. Batra for their collaboration in the studies of focused electron emission and scanning tunneling microscopy of Al, respectively. I also appreciate discussions with the faculty and research assistants of Department of Physics, Bilkent University in the course of this study, and comments by the examining committee.

My sincere thanks are due to my parents for their understanding and stimulating attention. Last but not the least, I owe a lot to my brothers Okan and Gürkan, and to Cemal Akyel for their continuous morale support.

Preface

The advents in lithography techniques and growth of high mobility samples motivated the studies on very small sized devices and structures. The efforts for miniaturization of electronic circuits and the quest for observing new physical phenomena in condensed systems yield the formation of a new area of condensed matter physics, namely mesoscopics. In the last decade research studies in mesoscopics increased in both quantity and quality and nowadays mesoscopics is a candidate for becoming one of the main subjects of condensed matter physics.

An important milestone in mesoscopics is the observation of quantized resistance. Although originating from a very simple quantum mechanics principle, realization and use of quantum point contacts lead to immense experimental and theoretical outcomes. Another tool of experimental physics, the scanning tunneling microscope, on the other hand, is shown to exhibit important resemblance to a quantum point contact.

In this thesis the ballistic transport and tunneling in laterally confined systems are investigated from a theoretical point of view. In Chapter 1 a brief review of mesoscopics is given with special emphasis on quantum point contacts and scanning tunneling microscopy. Chapter 2 and Chapter 3 are devoted to theory of quantum point contacts and scanning tunneling microscopy, respectively, in the context of mesoscopics. A seemingly different topic, namely nonequilibrium quantum transport, is studied in Chapter 4 in order to form a connection with simple quantum mechanical results and more realistic systems with inelastic scattering. A brief summary of the results and concluding discussions are in Chapter 5.

Contents

List of Figures

Chapter 1

Introduction to Mesoscopics

1.1 A Review of Mesoscopic Physics: Experiments and Theory

In the last few decades it is clarified that one of the most important objectives of the electronics technology is to miniaturize the devices. This yields both higher operation speeds and less power consumption. In short, the motto is *"small is beautiful"*. The development of technology is so steep that, for example, the size of memory cells decrease exponentially in time and is expected to be as small as a few hundred atoms in twenty years time as analyzed by Keyes[1] and shown in Figure 1.1. Nevertheless, this miniaturization trend is opposed by some physical constraints in addition to the technological ones. *Small is beautiful* only if the device operates according to the expectations, that is, only if all the related entities scale properly with the size of the device. For instance, a smaller *ohmic contact* has to be an ohmic contact with smaller conductance and so on. In Figure 1.1 also shown is this scaling behavior as studied by Guéret and coworkers.[2] Clearly, the scaling is not perfect as expected from the Ohm's law and there is a large variety in resistances of macroscopically identical contacts. This is partially due to the technological constraints, specifically due to the grain size effects. However, this scaling also has an intrinsic limit of applicability

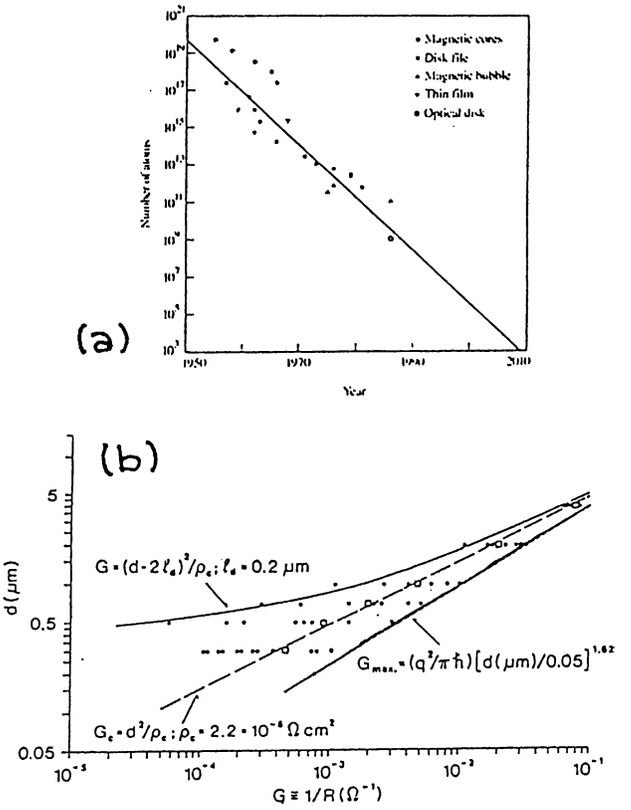

Figure 1.1: "Pros and cons" of miniaturization
(a) Number of atoms required to store one bit of information on a memory device. The straight line is a fitted exponential. From Reference 1. (b) Conductance of small ohmic contacts on GaAs as a function of contact size. The squares denote the average values obtained for macroscopically identical contacts. The full and dashed lines denote the extreme and average conductances, respectively. From Reference 2.

which becomes important when quantum mechanical and nonaveraging aspects of the system becomes dominant. This is only a single representative example for the complications of small systems from the point of view of technological and physical considerations. This study is devoted to the latter.

One of the most important features of the small systems is their sample specific properties. Our everyday experience conjectures that macroscopically identical systems have to yield the same results under identical experimental

conditions. However, for small systems this rule breaks down. As mentioned above, ohmic contacts fabricated on the same wafer using the same chemical and physical modification steps may have widely spread resistance values. Thinking in detail, this is very natural, however. It is known that the metal-semiconductor contact is not ordered and is made of grains. For a large contact, there is a large number of grains and the measured resistance is essentially an average of resistance of these grains. On the other hand, a small contact has only a small number of grains and this averaging can not be complete. Thus, the size and distribution of the grains may differ appreciably from one contact to other leading to the spread of resistance. Note that, this novel behavior is not solely due to the limited fabrication capabilities. Unless arranging the atoms one by one in a device, such effects will result independent of the technology used, due to the statistical nature of this arrangement. Taking a simpler example, for a macroscopic conductor contains a huge number of impurities and the resistance of the sample is determined according to Ohm's law or *microscopically* according to the Einstein or Drude formulae. For a small sample, however, there are only a few impurities and their strengths and positions determine the resistance. The averaging for such a small system is not sufficient to result in the diffusion relation of Einstein or the resistivity relation of Drude.

Another important aspect of small systems is the geometry-specific properties. Miniaturizing the devices further one reaches to a limit for which the device does not contain any impurities at all. For this case, the quantum mechanical propagation along the sample is essential. The material properties, however, are suppressed to a large extent. That is, the same results may be obtained by using completely different systems of the same geometry. This very effect is present for the case described in the preceding paragraph, as well. Although the material properties determine the mean behavior, the variation around this mean value is independent of the material and either *universal* or specific to the geometry. This is, in fact, is the manifestation of the *nonlocal* nature of transport in small systems. For instance, adding or removing a new geometrical feature to the circuit away from the *classical* path of the flow of the current the measured quantities

are changed.

At this point we wish to replace the term *small systems* with the term *"mesoscopic devices"* hereafter (as derived from the Greek prefix *meso*≡middle). The term mesoscopics corresponds to a length scale for which the averaging properties of the *macroscopic* world does not take place and the reversible and perfect mechanics of *microscopic* objects is not applicable. In what follows in this chapter we summarize the theoretical and experimental work has been done in the mesoscopic physics. The review is intended to be representative, but not exhaustive due to the limitation of space. More detailed information can be obtained from the recent conference proceedings on the subject.[3–6]

Aharonov-Bohm Effect

One of the most important manifestations of the effects listed above came out of the studies on Aharonov-Bohm effect in solid state devices. Aharonov-Bohm effect, or more precisely the *gedanken* experiment of Aharonov and Bohm,[7,8] is one of the most important demonstrations of the importance of the phase of quantum mechanical wave function. If a flow of current is divided into two branches as shown in Figure 1.2 and these branches enclose a region of nonzero magnetic field, the electron wave function will have differing phases depending on the branch it follows. Therefore, two components of current will interfere at the output leading to the so called Aharonov-Bohm oscillations. The phase difference between the electron wave functions going from the upper and lower branches is

$$\Delta\phi = \frac{e}{\hbar}\oint \mathbf{dl}\cdot\mathbf{A}, \tag{1.1}$$

A being the vector potential. This phase yields to oscillations in conductance of the circuit since the wave function squared at the exit is given by

$$|\psi|^2 \sim |\psi_1 + \psi_2 \exp(i\Delta\phi)|^2, \tag{1.2}$$

ψ_1 and ψ_2 being the wave functions corresponding to propagations through the upper and lower branches, respectively, for zero-field. Consequently, conductance (or resistance) exhibits oscillations periodic in the magnetic flux enclosed by the

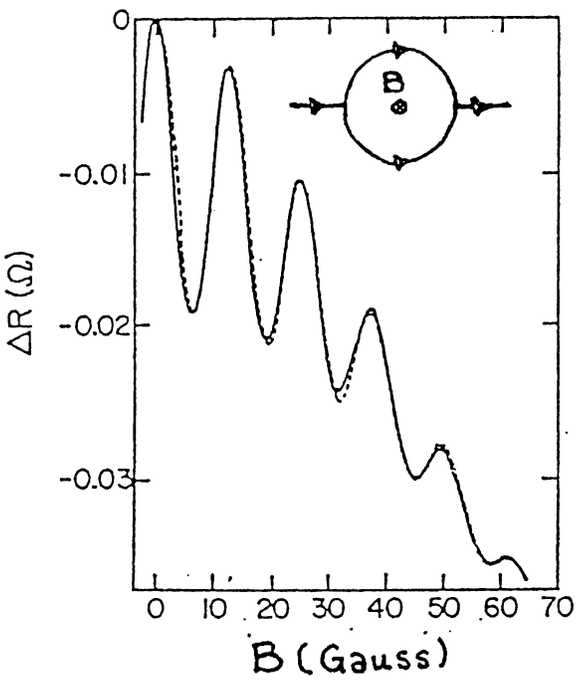

Figure 1.2: Aharonov-Bohm effect for a metallic loop
Change in the resistance of a lithium film evaporated onto a quartz fiber of diameter 1.3 μm and length 1cm. The magnetic field is parallel to the fiber. Temperature is 1.5 °K and the resistance at zero magnetic field is 2 KΩ. From Reference 9. The inset shows the schematics of the Aharonov-Bohm effect.

loop with period $\Phi_o = h/e$, the quantum of magnetic flux. The first experiments using solid state devices carried out by Sharvin and Sharvin[9] and Al'tshuler and coworkers[10] and yield the results shown in Figure 1.2. It took a few years until the western experimentalists reproduce these results. The magnetoconductance curve shown in Figure 1.2 contains a number of very important features as long as mesoscopics is concerned.

Strikingly, the period of oscillations is $\Phi_o/2$, and not Φ_o as expected. This point was clarified by Al'tshuler and coworkers.[10,11] According to their explanation, the $\Phi_o/2$ oscillations arise due to the interference of electrons enclosing the cylinder once clockwise and counterclockwise. Then the phase difference is twice of that given in Equation 1.1 and thus, the period halves. Along the loop electrons may be scattered by some impurities, yet scattering along the

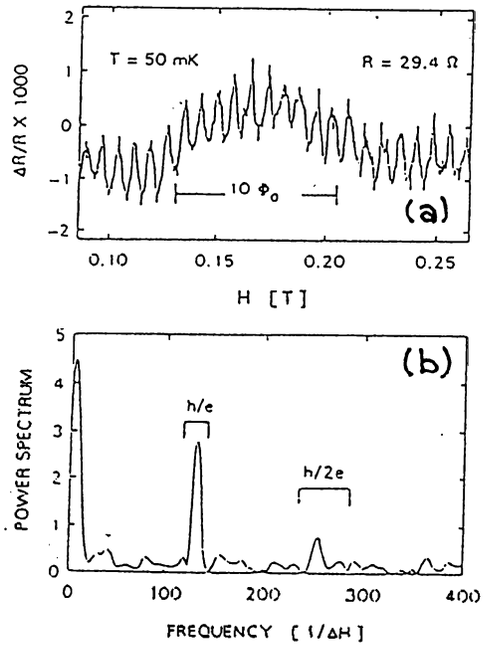

Figure 1.3: Aharonov-Bohm effect for a metallic ring
(a) Change in the resistance of a polycrystaline gold ring. The ring has inner diameter 0.8 μm, width 0.04 μm and thickness 0.04 μm. (b) The Fourier transform of the same data. From Reference 12.

two paths contribute exactly the same phase since they are time reversed forms of each other. On the other hand, the phase difference due to the scattering along the loop is different for electrons traversing only half of the circumference. Therefore, the Φ_o oscillations are phase averaged due to the randomly varying phase along the axis of cylinder. Using rings instead of cylinders it has become possible[12,13] to observe Φ_o oscillations together with $\Phi_o/2$ oscillations as shown in Figure 1.3. The phase averaging effect was also observed for N rings connected in series[14] for which the $\Phi_o/2$ oscillations remain constant in amplitude and Φ_o oscillations decay as $N^{-1/2}$ due to the random phase of Φ_o oscillations for each ring. Another important observation[9,10] in Figure 1.2 is the decay of the oscillation amplitude with increasing magnetic field. This is caused by the finite thickness of the cylinder, thus different trajectories contribute different Aharonov-Bohm periods which are out of phase.

The most important observation from a mesoscopics point of view is the diminishing oscillation amplitude with increasing temperature. This point, again, was explained by Al'tshuler and coworkers[10,11] using phase coherence arguments. Clearly, in a solid an electron suffers different kind of scattering processes. Some of these scattering events causes the electron to loose its phase. That is, knowing the phase of the electron wave function prior to the scattering one can not determine its phase after the scattering. This *incoherence* prevents the electron from displaying interference effects. It is possible, at least in principle, to define a phase coherence length L_ϕ as the distance over which the electron wave function retains its phase. If the circumference of the loop exceeds L_ϕ, then clearly the oscillations gets weaker. Al'tshuler and coworkers[11] showed that this decrease is exponential, $\Delta R_{AB} \propto \exp(-C/L_\phi)$, C being the circumference. In turn, L_ϕ decreases with increasing temperature due to the increased rate of inelastic scattering. This way, we have a criterion for the mesoscopic nature of any device. The related length scales are the phase coherence length L_ϕ, the electron mean free path ℓ_e, which is equal to the distance along which an electron moves without suffering any scattering, the Fermi wavelength λ_F, and finally the size of the device L. It is common to use the elastic scattering length instead of ℓ_e and the inelastic scattering length instead of L_ϕ. However, these substitutions may not be correct in general and one has to be careful when using them. For macroscopic devices one has $\lambda_F,\ \ell_e,\ L_\phi \ll L$ and therefore interference effects are not observable. For mesoscopic device $L \lesssim L_\phi$ and the interference effects are nor averaged out. The mesoscopic devices may have both $\lambda_F,\ \ell_e < L$ and $\lambda_F \approx L < \ell_e$. They are called to be in the diffusive[‡] and ballistic regimes, respectively. For the second class there are no impurities in the devices so that the geometry specific features are pronounced.

The last observation from Figure 1.2 is the negative overall magnetoresistance. This again is related to the interference of the states enclosing the flux once in opposite directions. This corresponds to a constructive interference for scattering

[‡]The term diffusive has to be differentiated from the usual diffusive transport, for which $L_\phi < L$. This regime is also referred as elastic diffusive regime.

from +k state to the −k state,[15] that is to a coherent backscattering. This yields a larger amplitude for −k state than would be expected and thus to a larger resistance for zero magnetic field. For a finite magnetic field, however, different loops contribute different periods and thus a phase averaging leads to the negative magnetoresistance. This point may be helpful to study the phase coherence length since as L_ϕ increases the effect of coherent backscattering gets stronger. For further interest on the solid-state Aharonov-Bohm effect we refer to existing review papers on the subject.[16,17]

Landauer Formulae

To this point the theoretical approaches to transport in mesoscopic devices have not been mentioned. Nevertheless, a complete understanding the mesoscopic physics requires a theoretical tool that can be used to find the conductance of a small sample for given arbitrary measurement conditions. For large samples phenomenological relations or approximations provide this tool. However, as described above for smaller samples these relations become irrelevant and a fully quantum mechanical theory which takes care of the position and strength of scatterers explicitly is required. Such approaches have been utilized since late 1950s. Kubo[18] and Greenwood[19] argued that the current through the sample is a cumulation of the local linear response of the system to an electromagnetic field. This electromagnetic field, in turn, is the external field plus the self-consistent field generated in response to it. The central entity in the Kubo formalism is the local conductivity tensor $\sigma_{\alpha\beta}$ and the local current density j_α is found by using

$$j_\alpha(\mathbf{r}) = \int d\mathbf{r}' \sigma_{\alpha\beta}(\mathbf{r},\mathbf{r}') E_\beta(\mathbf{r}'), \tag{1.3}$$

where E_β is the local electric field and α, β represent the cartesian components. In a transport measurement the local electric field may not be known due to the fact that the system is not homogeneous. Therefore, first the electric field $E(\mathbf{r})$ due to an applied bias has to be found, followed by the calculation of current density and the total current passing through the device. Although the Kubo formalism is proved to be appropriate for studying infinite systems, for mesoscopics it has

certain complications. In this regime it becomes necessary to calculate $\sigma_{\alpha\beta}$ for each individual device. In addition to that, in the mesoscopic regime not the local quantities as $\sigma(\mathbf{r}, \mathbf{r}')$ but the global quantities as conductance describe the device. Nevertheless, it was shown that linear response theory may be used in mesoscopics by taking proper cautions. The generalized Kubo formalism as developed by Středa and coworkers[20,21] is even more appealing to some theorists than the scattering approach described below.

In the same year as Kubo published his results, Landauer[22] used an alternative line of thought to derive an alternative approach to quantum transport. His motivation was the reciprocity principle of the electrical circuit theory. According to this principle, both the voltage across or the current through the circuit may be taken as the external driving agent. Then the response of the circuit is the current or voltage, respectively. Landauer argued that instead of applying an external electric field as it is done in Kubo formulation, one equally may draw a current through the device and gets exactly the same result for the conductance. This approach has the advantage over the standard Kubo formula that it is possible to include the boundary conditions easily and global quantities can be found without too much problem. Landauer's work and the following studies are widely referred as the *scattering approach* since they rely on the scattering of the incoming flux by the impurities in the sample.

The derivation of Landauer's formula for conductance for a strictly one-dimensional (1D) system is studied below in order to understand the underlying physics. Consider the system shown in Figure 1.4. The reservoirs on the left and right are assumed to satisfy the *blackbody* boundary conditions. That is, the incoming flux is totally absorbed and the outgoing flux is determined by the corresponding distribution function, for the case at hand by the Fermi-Dirac distribution. To this end, we assume that the left- and right-hand side reservoirs inject carriers into the *sample* up to Fermi levels μ_L and μ_R, respectively. The current injected by the left electrode is in excess of that injected by the right one. Ignoring the scatterer for the time being, the total current passing through the sample is simply the flux of electrons traversing the sample per unit time. This

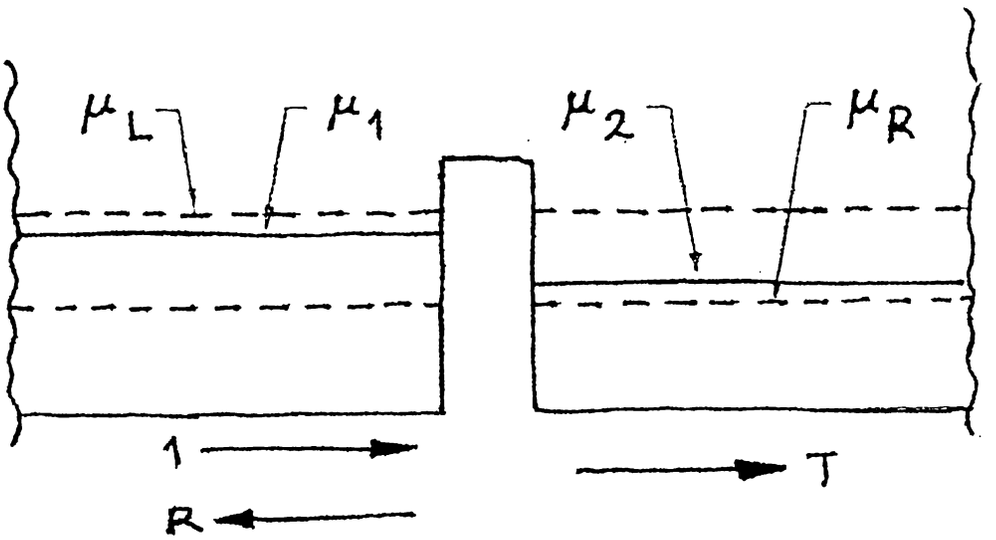

Figure 1.4: Calculation of conductance for a single scatterer

is equal to the product of the density of electrons and their velocity. The density, in turn, is the product of the density of states and the energy interval in which the left- and right-going electrons are unbalanced.

$$I = ev\mathcal{D}(E)(\mu_L - \mu_R). \tag{1.4}$$

where v is the velocity of electrons with energy E. The density of states is given by the usual expression

$$\mathcal{D}(E) = \frac{m}{\pi\hbar^2 k}, \tag{1.5}$$

where $\hbar^2k^2/2m = E$. Note that, the product $v\mathcal{D}(E)$ is equal to $1/\pi\hbar$ and is independent of energy. Using this relation the current is found as

$$I = \frac{2e}{h}(\mu_L - \mu_R). \tag{1.6}$$

In the presence of the scatterer only a portion of this current can pass through it and the rest is reflected back. Assuming that the transmission probability is T, the current becomes

$$I = \frac{2e}{h}T(\mu_L - \mu_R). \tag{1.7}$$

The difference of the Fermi energies, $(\mu_L - \mu_R)$, is chosen to be small so that the energy dependence of T within this range can be neglected. The voltage

difference across the scatterer, on the other hand, can be found by calculating the effect of the charge piled up near the scatterer. This can be done using either self-consistent screening[22] or the Einstein relation.[23] The electrochemical potential measured at the two sides of the scatterer can be found by employing a *counting argument.* The carrier densities on the left- and right-hand side of the scatterer are represented by μ_1 and μ_2, respectively. They can be determined by equating the number of occupied states above μ_i to the number of unoccupied states below μ_i for $i = 1,\ 2$. Noting that the total number of states between μ_R and μ_L is equal to $2\mathcal{D}(E)(\mu_L - \mu_R)$ and the density of electrons to the right and left of the scatterer are $T\mathcal{D}(E)$ and $(1+R)\mathcal{D}(E)$, respectively, one finds

$$T\ (\mu_L - \mu_2) = (2-T)\ (\mu_2 - \mu_R), \tag{1.8}$$

$$(1+R)\ (\mu_L - \mu_1) = (1-R)\ (\mu_1 - \mu_R), \tag{1.9}$$

$R = 1 - T$ being the reflection probability of the scatterer. The voltage difference

$$eV = \mu_1 - \mu_2, \tag{1.10}$$

can be easily found to yield

$$\mu_1 - \mu_2 = r(\mu_L - \mu_R). \tag{1.11}$$

Finally, the conductance $G = I/V$ which is given by

$$G = \frac{2e^2}{h}\ \frac{T}{R}. \tag{1.12}$$

Later Büttiker and coworkers[24] derived the multi-subband form of Equation 1.12 to find

$$G = \frac{2e^2}{h}\ \sum_n T_n \frac{2\sum_m v_m^{-1}}{\sum_m (1 + R_m - T_m) v_m^{-1}}, \tag{1.13}$$

where v_n, T_n and R_n are the velocity, total transmission and total reflection probabilities in the nth subband, respectively. The total transmission probability is obtained by adding the probabilities of transmission to the nth subband from all of the subbands $T_n = \sum_m T_{nm}$, T_{nm} being the transmission probability that

can be calculated by using the scattering matrix. Note that, the continuity of current for the multi-subband case requires $\sum_n R_n = \sum_n (1 - T_n)$.

The conductance formula Equation 1.12 had been rederived by several groups[25–28] by using various approaches. Anderson and coworkers[25] used this formula and a multi-subband generalization to study the scaling theory of localization. Engquist and Anderson[26] introduced the voltage probes in the system and found Equation 1.12 for weakly coupled voltage probes. Langreth and Abrahams[27] derived Equation 1.12 and its multi-subband generalization starting from the Kubo formalism. Azbel[28] was the first who derived Equation 1.13 using scattering states.

Synchronously studies giving different results for the conductance were appeared in the literature, as well. For example, Economou and Soukoulis[29] calculated the conductance for the strictly 1D case shown in Figure 1.4 using Kubo formalism and found

$$G = \frac{2e^2}{h} T, \tag{1.14}$$

as compared to Equation 1.12. Later, Fisher and Lee[30] employed the linear response theory for the multi-subband case and found that the conductance is given by

$$G = \frac{2e^2}{h} \mathrm{Tr}\{\tilde{\mathbf{t}}\tilde{\mathbf{t}}^\dagger\}, \tag{1.15}$$

$\tilde{\mathbf{t}}$ being the matrix of transmission probability amplitudes. The two sets of conductance formulae are approximately equal for the case $R \simeq 1$, that is, the opaque barrier limit. However, for the other extreme $R \simeq 0$, that is, for the transparent barrier limit Equations 1.12 and 1.13 yields infinite conductance, compared to the finite value obtained from Equations 1.14 and 1.15. These four formulae are usually referred as *Landauer formulae.* The first two are differentiated from the last two by classifying them as $G \sim T/(1 - T)$ and $G \sim T$ formulae, respectively. It is striking that different groups obtained different answers by using Kubo formalism. This difference is, in fact, due to the different assumptions made about the measured quantities, self-consistency etc.

The problem about the $G \sim T$ formula was that for a *perfect conductor* it gives a finite conductance $2e^2N/h$, N being the number of occupied subbands. This contrasts the intuitive expectation that the perfect conductor has no resistance. On the other hand, the $G \sim T/(1-T)$ formula gives infinite conductance and thus zero resistance for a perfect conductor. Imry[31] showed that the resistance given by the $G \sim T$ is a contact resistance and does not corresponds to the sample at hand, i.e, the perfect conductor. Using two containers of Fermions with different densities, and attached to each other by a narrow tube he found that the current through the tube is $I = \alpha(2e^2N/h)(\mu_L - \mu_R)$, α being a constant of the order of unity. In this example, the tube has no scattering center in it and perfectly matched to the containers so that it can be thought as a perfect conductor, and it still yields a finite conductance. This conductance is, in fact, associated with the contacts between the containers and the tube and is similar to the classical contact resistance. One may, at this point, think that for a classical transport measurement the effect of the contact resistances may be eliminated by using a four-probe geometry. That is, by using different probes for the current and voltage measurements. The resistance of a perfect conductor vanishes only for this four-probe measurement. Consequently, it has been thought that the $G \sim T$ formula applies for a two-probe measurement and $G \sim T/(1-T)$ for a four-probe measurement. This point had been taken for granted until the recent experimental results forced the theorists to think on the Landauer formulae once more as described below.

Universal Conductance Fluctuations

One of the most important problems arising from the results of the Aharonov-Bohm experiments carried out by rings[12,13] was that, in addition to the Φ_o and $\Phi_o/2$ periodic oscillations there were aperiodic fluctuations in resistance as a function of magnetic field as shown in Figure 1.3. These oscillations were reproducible for a single sample as long as it is not heated, but differ from sample to sample even though all the samples have the same macroscopical form. Thus, these oscillations were taken to be the *magnetofingerprint* of the

specific sample at hand. The period of these oscillations is larger than Φ_o and in the power spectrum as a function of frequency these oscillations appear as an increase near zero frequency. Strangely enough, same kind of oscillation were found for completely different mesoscopic devices, such as wires,[32] Si inversion layers[33] (conductance fluctuations as a function of the gate voltage) and for an Anderson model[34] (theoretical calculation of conductance). These systems have average conductances varying by orders of magnitude while the amplitude of the fluctuations is approximately constant and of the order of $\sim (e^2/h) \simeq 1/(25800\ \Omega)$ irrespective of the average conductance. Therefore, these fluctuations are often referred as the *universal conductance fluctuations*[‡].

The theoretical understanding of the universal conductance fluctuations followed the work done on *weak localization.* Up to this point we mentioned "disordered" systems or "metals" without explicitly defining these terms. It has long been recognized that for disordered systems some of the electronic states are localized and the transmission coefficient is exponentially small in the device size.[25] This situation is in contrast with the transport theories we discussed in the preceding subsection, since if the states near the Fermi level are localized it is not possible to talk about propagating electrons. Fortunately, the localization comes into the picture with a relevant length scale as the other physical effects do. This length is called the localization length ξ and the resistance of the sample increases as $\sim \exp(L/\xi)$ for $\xi < L$. This is generally referred as the strong localization regime. In this regime the transport occurs via Mott hopping or resonant tunneling. Thus, the density of states has spikes with large peak value but exponentially small resonance width.[35] Consequently the conductance has large fluctuations as a function of the Fermi level or impurity configuration. For strong localization regime the resistance is both non-Ohmic and have fluctuations which do not average out. These fluctuations are clearly not "universal" in the sense defined above.

Now consider the case for which $L < \xi$. Clearly the exponential decay

[‡]The "universality" of these fluctuations were discussed and defined with its limitations in the literature to which we refer in this subsection.

of the localized wave functions does not dominate the transport properties, rather the Bloch states are scattered by the disorder. Therefore, it is possible to use the Kubo or Landauer formalisms for quantum transport. This case is referred as the weak localization regime of metallic conduction. The early theoretical understanding on weak localization can be found in the related review papers.[15,36,37] It is possible to investigate the weakly localized sample in terms of the resonant tunneling transport as it is done for strong localization. For this case, however, the resonances are overlapping leading to a rather smooth density of states. Thus, it is possible to have fluctuations which are independent of the size of the system and the strength of disorder. In fact, within the weak localization theory it is shown that the density of states have important statistical aspects[38,39] which contribute to the conductance fluctuations.

Returning back to the universal conductance fluctuations, one observes that the transport in such weakly localized samples can be described by a random-walk type elastic diffusion. Therefore, one has a large number of randomly intersecting electron trajectories which interfere with each other as given by

$$t \sim \sum_{r,s} \exp\left[i(\phi_r - \phi_s)\right], \tag{1.16}$$

r and s denoting two different paths connecting the same end points and ϕ_r and ϕ_s are the respective phases attained going through these paths. Since this interference is not averaged out in mesoscopics and contributes to the conductance of the sample. Any perturbation that changes these phases causes fluctuation of the transmission probability, and thus of the conductance. Such a perturbation may be obtained by several means such as by changing the impurity configuration, by applying a magnetic field and thus adding an Aharonov-Bohm phase to the zero-field phase difference, and by varying the Fermi level. Although this qualitative explanation agrees with the experimental observations, the universality of the fluctuations can not be extracted therefrom.

The breakthrough for the quantitative understanding of the universal conductance fluctuations was the *ergodic hypothesis*.[40,41] According to this hypothesis the average of any statistical property of the conductance over

a sufficiently large magnetic field or energy range would be equal to the corresponding average taken for an ensemble of macroscopically equivalent systems. Therefore it suffices to consider the statistical properties of the conductance for an ensemble of impurity configurations. For example, the variance of these fluctuations is given by $< \Delta G^2 >$, $< \cdots >$ denoting the ensemble averaging and $\Delta G^2 = G^2 - < G^2 >$. This value was shown to be of the order of (e^2/h), independent of the average conductance of the system.[40–43] The exact number, however, depends on several details of the system and is not "universal". Another important quantity is the conductance correlation function $F(\delta B, \delta E)$ defined as

$$F(\delta B, \delta E) =< \Delta G(B, E)\Delta G(B + \delta B, E + \delta E) >, \qquad (1.17)$$

which determines the spacing of the fluctuations in magnetic field and Fermi energy. It was shown[40–43] that $F(\delta B, \delta E)$ is a decreasing function of its arguments with the corresponding half-widths B_c and E_c, respectively. The natural scale of magnetic field B_c is of the order of Φ_o/A, A being the area of the largest loop through which the phase coherence is retained. This is exactly the necessary magnetic field variation in order to change the phase difference of the interfering wave functions by 2π. The energy scale E_c, on the other hand, is of the order of the energy spread due to the inelastic scattering in the sample. $< \Delta G^2 >$, B_c and E_c can be calculated exactly as given by Lee and coworkers.[42] A systematic analysis of universal conductance fluctuations in Si inversion layers[44] showed that the agreement between the experiment and theory is excellent.

Multiprobe Measurements

The controversy on the Landauer formulae had been perceived as an academic problem until universal conductance fluctuations were observed. This was due to the fact that for most of the systems under consideration $R \simeq 0$ so that $G \sim T$ and $G \sim T/(1 - T)$ formulae give approximately the same answer. However, the universal conductance fluctuation experiments result an unexpected asymmetry in conductance.[32] This observation later supported by the magnetic field asymmetry of magnetoconductance of loops and wires[45]

and voltage fluctuations[46,47] measured for multi-probe devices. Clearly, the $G \sim T$ type Landauer formulae yields $G(B) = G(-B)$ and are not consistent with the experiments. The multi-subband generalization of $G \sim T/(1-T)$ type formulae as given in Equation 1.13, on the other hand, does not have such a symmetry. This asymmetry, however, is not sufficient to explain the experimentally observed asymmetry in universal conductance fluctuations.[34] Thus, the question concerning the relevant Landauer formula had become a serious confrontation with experimental results.

The symmetry of electrical conduction is a well known aspect of electromagnetic theory. The local conductivity tensor $\sigma_{\alpha\beta}$ satisfy the Onsager-Casimir symmetry relations[48]

$$\sigma_{\alpha\beta}(B) = \sigma_{\beta\alpha}(-B). \tag{1.18}$$

In addition, the reciprocity relation for multi-probe measurements were known since the first decade of the century in the absence magnetic field. It relies on the reciprocity of the current and voltage sources so that

$$R_{mn,kl} = R_{kl,mn}. \tag{1.19}$$

Here the multi-probe resistance $R_{mn,kl}$ is defined as the ratio of the voltage difference measured between the probes k and l to the current injected from probe m and drawn out of probe n

$$R_{mn,kl} = \frac{V_k - V_l}{I_{m\to n}}. \tag{1.20}$$

The generalization of the reciprocity relation to finite magnetic fields is more recent[49] and gives

$$R_{mn,kl}(B) = R_{kl,mn}(-B). \tag{1.21}$$

Although the presence of local symmetry as given by Equation 1.18 leads to the reciprocity relation, the reverse is not true. That is, it is possible to have a nonsymmetric local conductivity tensor and the global symmetry in Equation 1.21 is preserved.

Büttiker[50,51] argued that the discrepancy between the Landauer formulae and the experimental results arises from the handling of the voltage probes. In

deriving Equations 1.12 and 1.13 the voltage probes were taken to be weakly coupled to the device. This point earlier mentioned in the derivation of Engquist and Anderson.[26] On the other hand, Equations 1.14 and 1.15 correspond to a two-probe measurement. In the experiments, however, the lithographic shape of the current and voltage probes are the same. Thus, it is not possible to think about the weak coupling of the voltage probes. Büttiker considered the coherent device consisting of the probes in addition to the loop or wire, and calculated the scattering matrix for this *device*. All the probes are assumed to be connected to reservoirs in equilibrium. The current going into the device through probe m is found to be

$$I_m = \sum_m g_{mn}(\mu_m - \mu_n), \tag{1.22}$$

where g_{mn} is the generalized conductance for probes m and n, and is given by

$$g_{mn} = \frac{2e^2}{h} T_{mn}, \tag{1.23}$$

T representing the matrix $(\tilde{t}\tilde{t}^{\dagger})$. The electrochemical potentials of the probes $\{\mu_n\}$ are calculated by considering the current through the macroscopic connections of the probes. For example, in order to calculate $R_{mn,kl}$ one takes $I_m = -I_n = I$ and $I_i = 0$ for all the other probes. Inverting the matrix equation Equation 1.22, one gets all the electrochemical potential differences and thus, $(V_k - V_l) = (\mu_k - \mu_l)/e$. The four terminal resistance is $R_{mn,kl} = V/I$. Note that, this *multiprobe Landauer formula* is a natural extension of the original Landauer formulae, as far as the external excitation and scattering matrix are concerned. Büttiker[50,51] showed also that the four-probe resistance calculated from Equation 1.22 and 1.23 satisfies the global reciprocity relation Equation 1.21. The asymmetry of the magnetoconductance curves[45] was shown to be exact agreement with the predictions of this multi-probe Landauer formula.

It is interesting to study the limits of Equation 1.22 and 1.23. For the two-terminal resistance it reduces to Equations 1.14 and 1.15 as expected. For four-probes, voltage probes being weakly coupled, the $G \sim T/(1-T)$ formula is obtained for the single-subband case. For multi-subband systems the expression is complicated and does not simplify any conventional Landauer formula. The

multi-channel generalizations of $G \sim T/(1-T)$ are irrelevant due to the incomplete treatment of the voltage probes. It is also possible to obtain negative four-probe conductance as shown by Büttiker[51] for some systems. This is not contradicting any physical principle, since the power disspated in the circuit is proportional to the conductance found by using $G \sim T$ formula and is always positive. This formalism was also used to investigate the voltage fluctuations[52–55] due to the nonlocal nature of mesoscopic transport and had considerable success.

Finally, we wish to mention important contribution by Stone and coworkers[56,57] about the equivalence of the Kubo and Landauer-Büttiker formalisms. Starting from the standard linear-response theory and explicitly including the probes in their formalism they showed that the Kubo formulation yields Landauer-Büttiker relations both in the absence[56] and in the presence[57] of magnetic field. These derivations together with the success of Landauer-Büttiker approach seemingly ceased the controversy on the Landauer formula. We focus our attention on this discussion once more in the last section of this chapter.

1.2 Quantum Point Contacts

One of the most important breakthroughs in mesoscopic physics is the observation of the quantization of conductance in quantum point contacts (QPC).[58,59] Earlier different methods used to obtain mesoscopic devices in the diffusive regime as exemplified in Section 1.1. Note that, however, in the diffusive regime the quantum size effects are not essential due to relatively large device size and presence of scatterers. In the other regime of mesoscopics, namely for ballistic transport, the electrons traverse the device without suffering any scattering. Consequently, the size effects and geometry-specific features are the determining factors for ballistic devices. The fabrication of the ballistic devices was hindered due to the technological constraints until recently. In order to satisfy the requirements of ballistic transport the size of the device L has to be smaller than the electron mean free path ℓ_e and comparable to the Fermi wavelength λ_F. The first requirement guarantees that scattering in the device to be prohibited.

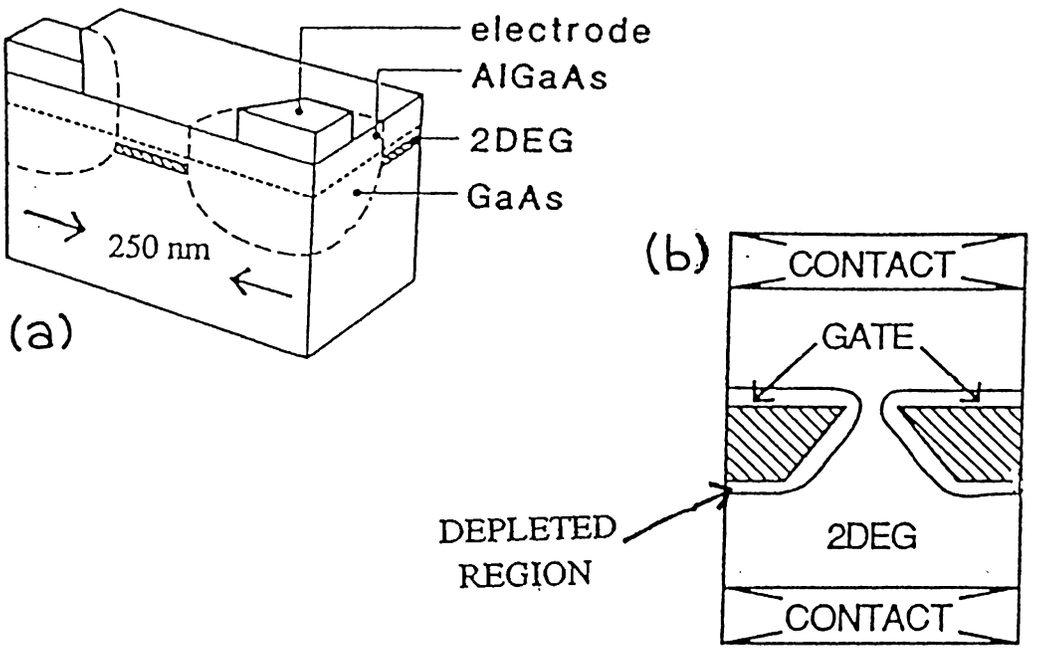

Figure 1.5: Schematic layout of a quantum point contact
(a) Cross-sectional and (b) top views of a quantum point contact

It can be achieved by using highly pure samples. The latter one, on the other hand, is the condition for the quantum size effects to come into the play. One way of achieving this is to decrease the electron density and thus to increase λ_F. However, for low density 2D electron gas (EG) the screening is weak and localization takes place. That is, it is necessary to miniaturize the device without decreasing the density. Therefore, the fabrication of QPC became a possibility only after the growth of high-mobility samples and advents in ion- and electron-beam lithography.

In Figure 1.5 the schematic layout of a QPC is shown. The 2D EG is obtained by a modulation doping AlGaAs/GaAs heterojunction and ought to have a mobility on the order of 10^6 cm^2-V^{-1}-sec^{-1} and the Fermi wavelength is a few hundred angstroms. On top of the sample a split-gate is defined by ion- or electron-beam lithography. The gap between the gates has to be comparable with λ_F and the width of the gate has to be smaller than ℓ_e. That is, both the width

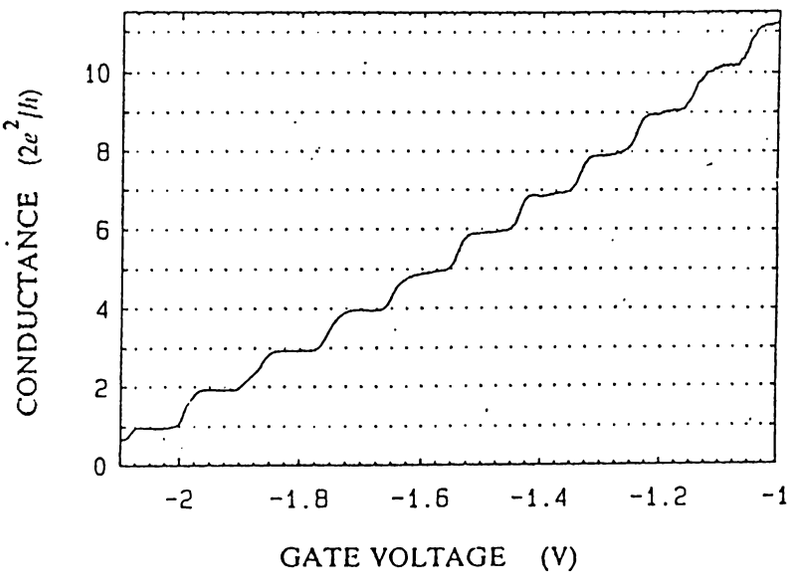

Figure 1.6: Quantization of conductance of a quantum point contact
The two-probe conductance of a quantum point contact as a function of the gate voltage. The temperature is 0.6 °K and the conductance is obtained using the data for resistance measurement after subtraction of background resistance of 400 Ω. From Reference 58.

and gap has to be a few hundred angstroms. The measurements are carried out via probes attached to the 2D EG by ohmic contacts and a negative gate voltage V_g is applied between the split-gate and substrate. For slightly negative gate voltages, the effect of the split-gate is small and the measurements yield results characteristic to 2D EG. For more negative gate voltages the electrons beneath the split-gate are depleted and a channel is obtained in the 2D EG. The characteristics of this novel device is quite different than that of a 2D EG since the electrons are *constricted* to flow from this channel. Increasing the negative gate voltage, the channel gets narrower due to stronger depletion. That is, it is possible to vary the width of the QPC externally for a given device configuration only by changing V_g. Further increasing the negative gate voltage all the electrons underneath the gate region are depleted, leading to the complete pinching off of the device. The easiest measurement that can be carried out by such a QPC is the measurement of the two-terminal resistance. The Delft-Philips collaboration[58] and Cambridge

group[59] succeeded to perform this measurement in late 1987. Their results were striking and the conductance (or resistance) of the QPC is found to change in steps for continuously varying V_g as shown in Figure 1.6.

In fact, the concept of point contacts was familiar to condensed matter physics for a while,[60,61] though, in a different context. The point contacts had been used to study the Fermi surfaces[60] and phonon scattering[61] in metals. The difference between the QPC and the earlier point contacts was the ratio of the device size to the Fermi wavelength, L/λ_F. As a result, only for the QPC quantum size effects are possible to play an essential role. For the earlier point contacts this ratio was of the order of hundred, compared to unity for QPC. Methods of classical electrodynamics and kinetic theory of gasses were used to analyze the conductance of point contacts. First Sharvin[60] incorporating the Fermion nature of electrons and using the Drude approximation for resistivity conjectured that the conductance of a point contact is independent of material parameters and is determined solely by the electron density and geometry. For a 2D point contact of width w, the Sharvin conductance is given by

$$G_s = \frac{2e^2}{h}\frac{2w}{\lambda_F}, \tag{1.24}$$

including spin degeneracy. Although the experimental results shown in Figure 1.6 are in qualitative agreement with the trend $G \sim w/\lambda_F$, the staircase structure is not consistent with the continuous change of G_s as a function of width w. The striking feature in the experimental results is that the change in the conductance between consecutive *plateaus* (where the conductance G is approximately stays constant with changing V_g) is equal to $2e^2/h$ within a precision of few percent. This *magic number* is the prefactor in Equation 1.24 and is called the quantum of conductance.

The quantum of conductance $2e^2/h$ was one of the central entities in mesoscopic physics as shown in Section 1.1. As early as 1957, Landauer[22,23] pointed out the essential connection between electronic transport and this magic number as described above. Later, Imry[31] showed that for a perfectly conducting narrow connection between two wide electron reservoirs the conductance is

of the order of this magic number. However, the observed quantization of the conductance in QPC had not been anticipated by theorists prior to the experiments. The reason of this can even be sought in the experimental results. To date, the precision of quantization in QPC is one percent, part of which may be thought as an intrinsic deviation from quantization[‡]. On the other hand, e^2/h can be measured by use of quantum Hall effect[62] within an accuracy of one part in 10^7. That is, the *quantization of conductance* in a QPC is not exact and, in fact, highly distorted. Before the pioneering experiments[58,59] this distortion was expected to wash out the main effect, namely quantization. Actually, *dirty* electron waveguides, which does not yield quantization, were used to understand the diffusive regime of mesoscopics as depicted in the preceding section. Only after the experiments it became clear that the quantization may be robust against several interfering effects.

We have to admit that the development of the nanotechnology took over the theoretical studies and experiments leaded the theorists in the subject of QPC at the early stages. Nowadays, the QPC is used to analyze several important aspects of electronic properties and transport in condensed systems. Especially, magnetotransport measurements carried out by use of QPC provide unique opportunities to investigate the behavior of mesoscopic systems with applied magnetic field. These subjects, however, are beyond the scope of this study. To our knowledge a wide review of the studies of the Delft-Philips collaboration constitutes the best reference for the experimental work that have been done in the field.[63] On the other hand, several groups shed light on the theoretical basis of the quantization of conductance as described in Section 2.1. The current focus of interest from theoretical point of view is the reexamination of Landauer formulae and electronic transport in mesoscopic systems by considering QPC and electron waveguides as the main building blocks.

The existing theoretical tools at the time of the experiments were sufficient to explain the quantization in simple terms.[58,59] The subband formation was well understood and exploited earlier.[64] Clearly, the accumulation layer of the 2D EG

[‡]This point is later verified by the theoretical studies as shown in Chapter 2.

and the depletion due to the negative gate voltage prevents the electrons from escaping out of the channel. On the other hand, the electrons are free to move along the channel. This situation is reminiscent of a simple textbook problem for an *electron waveguide* and it can be shown that the lateral momentum of the electron is quantized yielding subband formation. The subband energies can be expressed within effective mass approximation as

$$E(n_x, n_y, k_z) = \epsilon_{n_x n_y} + \frac{\hbar^2 k_z^2}{2m}, \tag{1.25}$$

where n_x and n_y are the quantum numbers associated with the quantization of the components of the momentum perpendicular to 2D EG and perpendicular to the channel in the plane of the 2D EG, respectively, and k_z is the component of the momentum along the channel. $\epsilon_{n_x n_y}$ is the eigenenergy as found from the 2D particle in the box problem and for an infinite-well box is given by

$$\epsilon_{n_x n_y} = \frac{\hbar^2}{2m}\left[\left(\frac{\pi}{n_x L_x}\right)^2 + \left(\frac{\pi}{n_y L_y}\right)^2\right], \tag{1.26}$$

L_x and L_y being the widths of the accumulation layer and channel, respectively. Equation 1.25 implies that propagation is allowed only for subbands with $\epsilon_{n_x n_y} < E_F$, E_F being the Fermi energy. For the QPC under consideration the width of the accumulation layer is so thin ($L_x \sim 100$ Å) that only the lowest subband of the 2D EG, i.e., $n_x = 1$, is occupied. Therefore, we neglect the dimension perpendicular to the 2D EG completely and use a strictly 2D system from now on by suppressing the index n_x. This is a good approximation as long as the properties related directly to the 2D EG, such as the effects of the surface roughness, are not considered explicitly and is used throughout Chapter 2. Accordingly, only N_p subbands with $\epsilon_{N_p} < E_F < \epsilon_{(N_p+1)}$ are propagating in the channel.

Assuming that this channel is attached to two reservoirs at $z = \pm\infty$ kept at constant electrochemical potential with an infinitesimal difference $\Delta\mu$ it is possible to find the conductance of the QPC. The current is given by a multi-subband generalization of Equation 1.4

$$I = \sum_{n_y} e v_{n_y} n_{n_y} = \sum_{n_y} e v_{n_y} \mathcal{D}_{n_y}(E_F) \Delta\mu, \tag{1.27}$$

where v_{n_y} is the velocity and $\mathcal{D}_{n_y}(E)$ is the density of states for electrons in subband n_y. The cancellation of the velocity from this expression is valid for the multi-subband case as well since

$$\mathcal{D}_{n_y}(E) = \theta(E - \epsilon_{n_y}) \frac{m}{\pi \hbar^2 k_{n_y}}, \tag{1.28}$$

with $\hbar^2 k_{n_Y}^2 / 2m = E - \epsilon_{n_y}$. Thus the current is found to be

$$I = \sum_{n_y} \theta(E_F - \epsilon_{n_y}) \frac{e}{\pi \hbar} \Delta\mu. \tag{1.29}$$

The conductance of the QPC in a two terminal configuration is calculated by using the difference of the electrochemical potentials of the reservoirs as the voltage drop along the circuit, i.e., $\Delta\mu = eV$. Therefore the two terminal conductance $G = I/V$ is quantized

$$G_p = \frac{2e^2}{h} N_p. \tag{1.30}$$

This expression can also be obtained in terms of the Landauer formula $G \sim T$ by using the identity matrix as the transmission matrix and corresponds to the *conductance* of a perfect conductor. Thus, the main difference between this simple model and the real structure is that in the latter the 2D EG on two sides of the point contact were taken to be ideal reservoirs and consequently the propagation through the subbands are independent of each other. This assumption may not be valid as depicted by the poorness of the quantization in the experiments, however. In addition, the length of the channel in the experiment is of the order of λ_F and is not infinite as this simple explanation requires. Nevertheless, this first explanation[58,59] incorporated most of the essential aspects of quantization of conductance. The remaining theoretical problem was to investigate the properties of the quantization and effects of imperfections.

The subject matter of Chapter 2 is the analysis of quantization of conductance in a QPC. Being motivated by the above simple argumentation we give emphasis to the quasi-1D nature of transport in the constriction. That is, we use the subband structure explicitly in our formalism. The essence of our method is to match the wave function in the channel to that in the 2D EG *reservoirs*.

This way, the conductance of the QPC is expressed in terms of the coefficients of the subbands in the total wave function, in conjunction with the above simple explanation. Treating the 2D EG as a free electron gas, a new feature, namely the interference of the waves in the channel, comes into the play due to nonzero reflection probability from the ends of the channel. This aspect yields a resonance structure to be superimposed on the quantized plateaus which is not observed in the experiments. Therefore, we focus our attention to this resonance structure and analyze its properties for different QPC systems. It is found that for long constrictions the thermal averaging may wash out the resonance structure without affecting the quantization. Still, this can not explain all the experiments especially those carried out in extremely low temperatures. In the course of this analysis we find that the most important advantage of the present approach is the analogy to the strictly 1D problems, and the versatility resulted therefrom. In this context, we adapt the transfer matrix method in order to handle problems concerning nonuniform constrictions. This enables us to consider several QPC geometries without making exhaustive computations. The effects of rounding the opening of the constriction into the 2D EG is analyzed and it is found that the adiabatic evolution of states, i.e., the suppression of reflection in the channel, eliminates the resonance structure. The distortion of quantization due to elastic scattering by impurities in the channel is investigated again by exploiting the analogy between the QPC and strictly 1D systems. The enhancement of backscattering due to the states localized in the vicinity of the attractive impurities is conjectured. Resonant tunneling via the quasi-bound states are analyzed and the forementioned analogy is verified.

1.3 Scanning Tunneling Microscopy and Mesoscopics

One of the essential requirements for mesoscopic behavior is that the quantum mechanical wave function and especially its phase determine the observable

physical properties of the system. As summarized in Section 1.1 systems having this property had been fabricated only recently. Another quantum mechanical *device* developed in the last decade is the revolutionary scanning tunneling microscope (STM).[65,66] That STM satisfies this very property, however, has gone unnoticed for half a decade.[‡]

An STM essentially consists of an atomically sharp metallic tip and a sample (metal, semimetal or semiconductor) to be investigated. The tip and sample are brought into vicinity of each other with a potential barrier in between (vacuum, air or liquid). A schematic description is given in Figure 1.7. Applying a potential bias between them a tunneling current passes through the circuit. This tunneling current is a function of the lateral position of the tip in addition to the tip-sample distance, due to the atomic size of the tip. Thus, *scanning* the sample by moving the tip around by using a piezoelectric positioner, one gets information about the topographical and electronic structure of the sample. Several interesting applications[66] of STM have been used since the invention of the apparatus by Binnig and Rohrer.[65] However, the main usage of STM had been the characterization and analysis of surface structure for the first few years.

The appealing surface imaging capabilities of STM had attracted interest from theoretical point of view as well. The basics of the operation of STM was the well known theory of tunneling. Nevertheless, the superior resolution obtained by STM is unexpected and the exact interpretation of the results were tricky. Shortly after the first results obtained by STM were reported, Tersoff and Hamann[67] provided an elegant theory for the *device*. They made use of the transfer Hamiltonian method as described by Bardeen[68] and modeled the tip by a spherical potential well. At the end they found that the conductance of the vacuum barrier is given by

$$G_{TH} = \frac{2e^2}{h} \{32\pi^4\varphi^2\mathcal{D}_t(E_F)R^2\kappa^{-4}e^{2\kappa R}\}\ \rho_s(\mathbf{r}_o; E_F), \tag{1.31}$$

where φ is the height of the tunneling barrier, $\mathcal{D}_t$ is the density of states of the tip, R is the radius of the tip and $\kappa = \sqrt{2m\varphi}/\hbar$ is the inverse decay length of

[‡]The designation STM is used to denote both "the scanning tunneling microscope" and "scanning tunneling microscopy" throughout the text.

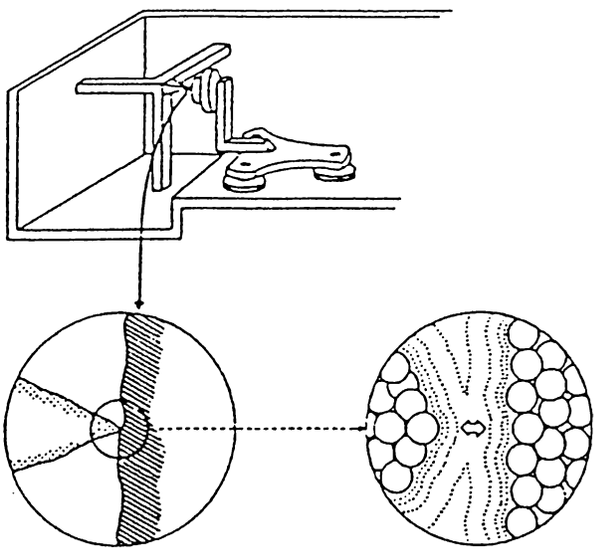

Figure 1.7: Schematics of the scanning tunneling microscope
The mechanical parts include three piezo-drives, the tip, sample holder and the louse. The tip-sample region is amplified to atomic scale.

wave functions in the potential barrier. The central quantity, namely $\rho_s(\mathbf{r}_o; E_F)$ is the local density of sample states at the center of the tip $\mathbf{r}_o$, evaluated at the Fermi energy. It is not surprising to have $(2e^2/h)$ as the prefactor since it has the dimension of conductivity and expressed in terms of fundamental physical constants. Equation 1.31 is, in fact, nothing but the Landauer formula $G \sim T$ for an exotic geometry. The beauty of Tersoff-Hamann theory lies in the fact that if certain conditions are satisfied,[67] it is not necessary to solve the transmission problem for this complicated geometry. Instead, it is possible to express the conductance as a product of tip-related quantities (the term in parenthesis in Equation 1.31) and sample-related quantities (the local density of states). This is a result of the special tip shape used by Tersoff and Hamann. Nevertheless, this assumption was verified to a certain extent by the experimental results[66] as well.

Equation 1.31 is different from what we have seen for the QPC in the preceding section in such a respect that it does not lead to any kind of quantization of

conductance. Moreover, the subband structure, if ever exists, is not manifested by this expression. Thus, one has to conclude that STM is intrinsically a 3D device and its connection to mesoscopics is not more than that of any small device. This statement is incomplete, if not wrong. As mentioned above, Equation 1.31 is valid only if some conditions are satisfied. These conditions are, in fact, the validity conditions for the transfer Hamiltonian explanation of tunneling.[68] Clearly, when the transmission probability is large or when the electrodes are very close to each other, these conditions are not satisfied. Therefore, the small tip-sample distance operation of the STM has to be analyzed in detail before concluding whether STM is a mesoscopic device or not[‡]. For example, pushing the tip towards the sample until they touch each other it should be possible to obtain a point contact. This point contact may be in the quantum regime as well as Sharvin regime, depending on the size of the contact. For such a point contact it is natural to use the formulation developed in Chapter 2. Thus, one expects to have some relation between the STM at small tip-sample distance operation and mesoscopics.

A striking experiment partially shed light on this point.[69,70] In this experiment Gimzewski and Möller studied the dependence of the tunneling gap resistance as a function of the tip-sample distance to understand the transition between the tunneling regime and point contact in STM. Their results are shown in Figure 1.8 and incorporate several important aspects of STM. First, $\log I$ does not change linearly with tip-sample distance. This may have several reasons, including the tip-sample interaction and the resulting reduction of the tunneling barrier height. Second, there is an abrupt jump in current and after this jump the conductance is of the order of $(2e^2/h)$. This abrupt jump is related to the mechanical contact of the tip to the sample and afterwards a point contact, possibly in the quantum regime, is formed between the tip and sample. Finally, for some experiments some oscillations had been observed further pushing the tip towards the sample after this mechanical contact. These oscillations may have some resemblance to the quantized conductance in the QPC.

[‡]The effects of the atomic scale tip-sample interaction on tunneling current had been investigated earlier, and a summary is given in Chapter 3. In this section we focus our attention on the connection between STM and quasi-1D transport only.

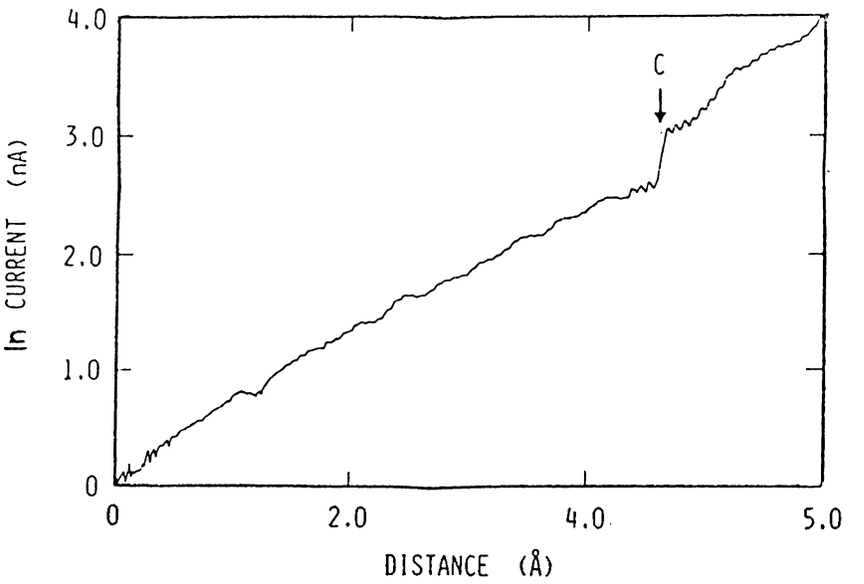

Figure 1.8: Transition from tunneling to point contact in scanning tunneling microscopy
Tunneling current as a function of distance for an Ir tip and polycrystaline Ag sample for a bias of -20 mV. From Reference 69.

All these features can be resolved by using the tip-sample interaction and quasi-1D transport in conjunction. This requires the adaptation of the formulation developed for the QPC to the STM geometry. This is the subject matter of Chapter 3. We present an alternative theory for STM which is capable of explaining the transition from tunneling to point contact and relying on the analogy to the QPC. Both the approach to point contact and the conductance after the abrupt jump are analyzed.

Although the present formalism is beneficial for understanding phenomena taking place close to the point contact, its applicability for the standard STM operation seems to be small. Nevertheless, we test it by considering one of the anomalous results obtained by STM. Recently, atomic corrugation on nominally flat noble[71] and simple[72] metal surfaces have been achieved. In Figure 1.9 the result obtained for Al(111) surface is shown. The measured corrugation is one order of magnitude larger than that predicted by Equation 1.31. We show that this anomalous corrugation can be explained considering the tip-sample

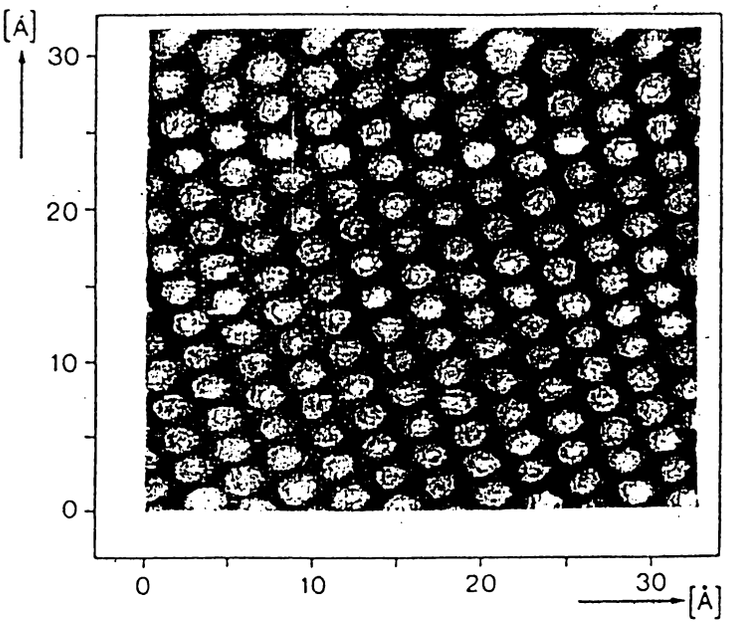

Figure 1.9: Atomic corrugation on Al(111) surface obtained from scanning tunneling microscopy
Grey-scale image of a 34 $\mathring{A} \times 34\mathring{A}$ area for a bias of -50 mV and tunneling current is kept constant to be 6 nA. The height corrugation is 0.3 $\mathring{A}$. From Reference 72.

interaction effects and using the quasi-1D channel model for STM. In Chapter 3 we finally investigate the focusing properties of electron emitting sharp tips, i.e., the so called *point sources*, and QPC. This topic have become popular after the fabrication of single atomic stable tips[73] and use of them in electron and ion emission experiments[74] by Fink. Our results are not only in qualitative agreement with the experimental results, but resolve the existing controversy about the effect responsible for focusing as well. We find that the effective barrier due to the lateral confinement of the current carrying states and the horn-like opening to the vacuum are causing the collimation of the emitted beam.

1.4 "Open Problems about Open Systems"

In a transport measurement the mesoscopic device is driven by external sources (current or voltage) and the response of the system to these excitations (usually

a current or voltage) is measured by an apparatus connected to the device. A correct quantum mechanical theory dealing with such a transport process, in principle, has to consider both the external sources and measurement apparatus quantum mechanically, that is, in the form of a relevant Hamiltonian. This is, however, beyond the capabilities of contemporary quantum transport theories. The most important reason for this is the macroscopic size of the external sources, measurement apparatus and their connection to the device. A simplification can be done by assuming that the system, which consists of the device only, is not closed. Open systems, however, have serious problems. The problems associated with open systems were first pointed out by Landauer[75] in the context of mesoscopics. For example, since in the transport process there is a flow of carriers along the device, the continuity equations are not satisfied. Secondly, for an open system the boundary conditions are not strict and may cause some arbitrariness in the solution. All such problems may be summarized by saying that an open system does not have a hermitian Hamiltonian. The everyday-quantum mechanics, on the other hand, has been formulated mostly for hermitian operators. Therefore, it is necessary to make some modification on the system in order to investigate it theoretically.

One way of circumventing these problems is to use a closed system, which acts as an open system to a certain extent. This puzzling statement can be illustrated by an example originally proposed by Imry.[31] Two very large containers with two different densities (or equivalently electrochemical potentials) of fermions may represent two disconnected reservoirs. They thus act as a quantum mechanical battery.[76] Connecting these two reservoirs to each other by using a mesoscopic device (for example, a QPC) particles will start to flow out of the reservoir of higher density to the other one. This flow is equivalent to a current through the mesoscopic device. If the volume of the containers is large enough and the particle density is low enough then the flow of particle will oscillate in time (just like that in a double-well structure). The period of this oscillation, however, will be very large due to the same reasons. Therefore, it is possible to make some transport measurements on this system as if it is an open one. This is possible only if the

time span of the measurement is very small compared to the oscillation period. However, this model has some unphysical features as well. For example, the process described above is completely reversible and no dissipation takes place. It is necessary to include some dissipation mechanism that will drive the system to an equilibrium, i.e., equal densities in both reservoirs. This way the system becomes irreversible, that is, theoretically one has to calculate the decay of a nonequilibrium state into an equilibrium one. Some kind of Boltzmann equation would form an appropriate basis for such an approach. Apparently, even though the system is closed one has to have a nonhermitian Hamiltonian in order to prevent reversibility of the system. The importance of inelastic scattering in this context has been discussed by Landauer.[75]

A second way of examining a transport measurement is to work with open systems. To this end, only the device under consideration is represented quantum mechanically in the Hamiltonian. The external source and measurement apparatus, on the other hand, are represented by fictitious reservoirs. Usually these reservoirs are taken to satisfy the blackbody boundary conditions described in Section 1.1. Although the two approaches seem alike[‡] they have some essential differences. For the first one, one has to design some kind of reservoir to simulate the circuit elements. For the latter, however, the internal structure of the reservoirs is irrelevant and only the boundary conditions (such as the current through or the potential at the probe) imposed by that reservoir takes place. The usual way of dealing with this problem is to formulate a scattering approach and to solve the scattering matrix for the given boundary conditions. This is, in fact, nothing but the multiprobe generalization of Landauer formula.[50,51] This simple explanation, however, is not straightforward as it stands. As mentioned in Section 1.1 a controversy which lasted more than one decade demanded a deeper understanding and more refined analysis of quantum transport in mesoscopic devices. Even nowadays it is hard to say that the controversy is completely over.[56,57,75,77] In what follows we summarize the arguments leading to this

[‡]One way of obtaining a blackbody is to have a large irregular cavity with a very narrow opening.

controversy and further problems on open systems.

Stone and Szafer[56] recently derived the multiprobe Landauer formula Equation 1.23 as proposed by Büttiker[50,51] from linear response theory. They also discussed the reasons for presence of numerous Landauer formulae which can also be derived as a linear response. The apparent difference between the approaches that came out with $G \sim T/(1-T)$ formulae[25–27] and $G \sim T$ formulae[29,30] is the inclusion of self-consistent field for the former ones. This inclusion relies on the argument that the electron-electron interaction has to be taken into account in order to prevent formation of charge density waves in the ideal conductors connecting the disordered device to the reservoirs. Stone and Szafer[56] argued that this charge density wave is a result of the boundary conditions imposed by the steady-state current driven by the external source and is a part of the linear response of the system. Landauer,[77] on the other hand, in a critical review of theoretical studies of quantization of conductance in QPC argued that the high space charge due to the unbalanced electron flow through the constriction yields a self-consistent field, which gives rise to corrections to the two-probe conductance result given by Equation 1.30. In fact, earlier Thouless[78] reported in a comment that the solution obtained by ignoring the presence of the charge density wave is $G \sim T$ formula, and the inclusion of the self-consistent field yields corrections precisely what is needed to get $G \sim T/(1-T)$ formula. Very recently Büttiker[79] gave credit to both approaches in a work reviewing the quantum transport and multiprobe measurements.

Both approaches to quantum transport,[56,77] however, raise interesting questions about the nature and importance of reservoirs and measurement probes. Although the measurability and relevance of electrostatic potential is still an open issue,[56,76,77] a quantum mechanical description of *phase-randomizing* reservoir and *non-invasive* voltage probes is a problem common to both groups. The importance and applicability of different Landauer formulae is another question which deserves further investigation. Finally, the effect of inelastic scattering has to be considered in conjunction with the above questions. In Chapter 4 we seek for answers to these ''open questions about open systems'' (as Stone

and Szafer[56] named them) by using the Keldysh technique for nonequilibrium Green's functions. Despite the model presented is not detailed enough to cope with realistic problems, we find results which may shed light on the continuing debates on quantum transport.

Chapter 2

Theory of Ballistic Transport through a Quantum Point Contact

2.1 Theory and General Formalism

As the simple explanation described in Section 1.2 implies the essential feature responsible for the quantization of conductance in the QPC is the quasi-1D nature of the system. For an infinitely long uniform channel the quantization is exact. However, the experiments[58,59] were carried out with QPC of finite length. Therefore, the quantization is distorted depending on the parameters of the geometry. The subject matter of this section is to devise a formalism which is applicable to various QPC geometries, and thus suitable for carrying out a systematic study on ballistic transport through a QPC.

In order to emphasize the similarities with the infinite uniform channel, we include the quasi-1D character of transport explicitly in our formalism. This is done by separating the space into three parts. The leftmost and rightmost parts are the 2D EG which are connected to the reservoirs. The EG are semiinfinite and 2D free electron wave functions are solutions of the Schrödinger equation. The central part is the constriction which is characterized by a laterally

confining potential. Consequently, the solutions of the Schrödinger equation in the constriction are the subband wave functions arising from the quantization of transverse momentum. We use the effective mass theory[‡]throughout this chapter. This approximation is verified to a great extent since the size of the systems we are considering are so large that they include thousands of atomic cells and in addition the structure is lightly doped.

The separation of the space into the 2D EG and constriction can be represented in the Hamiltonian by using potential $V(y,z)$ given by

$$V(y,z) = [V_s(z) + V_c(y,z)]\ \theta(z)\ \theta(d-z). \tag{2.1}$$

Here the potential is exactly zero in the left- ($z < 0$) and right-hand side ($z > d$) 2D EG. In the constriction ($0 < z < d$) the potential is decomposed into two components. The longitudinal part $V_s(z)$ includes the variation of the minimum value of the potential along the direction of propagation. The confining part $V_c(y,z)$, on the other hand, is responsible for the formation of the subband structure. For a general constriction potential $V(y,z)$ the decomposition into longitudinal and confining parts is not straightforward, and even may not be unique. To circumvent this difficulty in our analysis we rather start with the component potentials but not with the full potential. This way the roles and effects of $V_s(z)$ and $V_c(y,z)$ are clarified.

The Hamiltonian for the QPC can be written as

$$H = -\frac{\hbar^2}{2m^*}\nabla^2 + V(y,z), \tag{2.2}$$

which reduces to the kinetic term in the 2D EG. Here m^* is the effective mass for electrons propagating in the 2D EG, which is assumed to be isotropic. The subband wave functions are calculated by solving the Schrödinger equation in the

[‡]As known the effective mass theory treats the effect of the atoms as a renormalization of the free propagation energy. It is appropriate to use the effective mass approximation for systems with extended Fermi surfaces, containing a large number of atomic cells, and with external potentials varying smoothly over atomic length scales. All these conditions are satisfied for the system at hand.

constriction

$$\left[-\frac{\hbar^2}{2m^*}\frac{\partial^2}{\partial z^2}+V_s(z)-\frac{\hbar^2}{2m^*}\frac{\partial^2}{\partial y^2}+V_c(y,z)\right]\xi_{nE}(y,z)=E\xi_{nE}(y,z), \qquad (2.3)$$

where we assumed that the energy spectrum is continuous due to the propagation along the z-direction. Also note that due to this propagating nature of the wave function the states are two-fold degenerate, so that $\xi_{nE}(y,z)$ and $\bar{\xi}_{nE}(y,z)$ satisfy Equation 2.3 with left- and right-going current densities, respectively. Since the partial differential equation in Equation 2.3 is not separable in general, in principle it has to be solved as it stands. However, this is possible only for special potential profiles. In Section 2.3 an approximate way of solving this equation for a general potential profile is presented. For the time being the focus of attention is placed on the calculation of conductance given the subband wave functions $\xi_{nE}(y,z)$ and $\bar{\xi}_{nE}(y,z)$.

It is constructive to recall that the essential feature of the present model is the separation of the space into 2D EG and the constriction. Equation 2.3 represents the wave function in the constriction. However, in a transport measurement the electrons are originated from one 2D EG and collected into the other one. Therefore, the subband wave function described in Equation 2.3 has to be matched to the 2D plane waves in the EG. This is achieved by employing the continuity of the wave function and its derivative at the boundaries between the constriction and the left- ($z = 0$) and right-hand side ($z = d$) 2D EG. The boundary conditions at $z \to \pm\infty$, on the other hand, define the incoming and outgoing waves. In what follows we determine the wave function in the whole space for an incident plane wave from left-hand side 2D EG with wave vector $\mathrm{k}_i = (\kappa_o, k_o)$ and energy $E = \hbar^2|\mathrm{k}_i|^2/2m^*$.

For incidence from left, one has the incident wave and reflected waves in the left-hand side 2D EG and the outgoing waves in the right-hand 2D EG, which are linear combinations of plane waves. These plane waves and the subband wave functions in the constriction are of the same energy E, since the inelastic processes are not included. That is, the wave function $\psi_{\mathrm{k}_i}(y,z)$ can be written

as

$$\begin{aligned}
\psi_{\mathbf{k}_i}(y,z) &= e^{ik_o z}\, e^{i\kappa_o y} + \int d\kappa\; e^{-ik_z(\kappa)z}\, e^{i\kappa y}\, A_{\mathbf{k}_i}(\kappa), && z<0,\\
&= \sum_n \{\xi_{nE}(y,z)\Theta_{n\mathbf{k}_i} + \bar{\xi}_{nE}(y,z)\Delta_{n\mathbf{k}_i}\}, && 0<z<d,\\
&= \int d\kappa\; e^{ik_z(\kappa)z}\, e^{i\kappa y}\, B_{\mathbf{k}_i}(\kappa), && d<z, \qquad (2.4)
\end{aligned}$$

where $k_z^2(\kappa) = 2m^*E/\hbar^2 - \kappa^2$, imaginary part of $k_z(\kappa)$ being positive in order to have a finite wave function for $z \to \pm\infty$. The wave function evaluated at the boundaries of the constriction are

$$\begin{aligned}
\psi_{\mathbf{k}_i}(y,0) &= e^{i\kappa_o y} + \int d\kappa\; e^{i\kappa y}\, A_{\mathbf{k}_i}(\kappa)\\
&= \sum_n \{\xi_{nE}(y,0)\Theta_{n\mathbf{k}_i} + \bar{\xi}_{nE}(y,0)\Delta_{n\mathbf{k}_i}\}, \qquad (2.5)
\end{aligned}$$

and

$$\begin{aligned}
\psi_{\mathbf{k}_i}(y,d) &= \sum_n \{\xi_{nE}(y,d)\Theta_{n\mathbf{k}_i} + \bar{\xi}_{nE}(y,d)\Delta_{n\mathbf{k}_i}\},\\
&= \int d\kappa\; e^{ik_z(\kappa)d} e^{i\kappa y}\, B_{\mathbf{k}_i}(\kappa). \qquad (2.6)
\end{aligned}$$

The continuity of the derivative along y-direction is guaranteed by the continuity of the wave function itself. On the other hand the derivative of the wave function along z-direction evaluated at the boundaries of the constriction are

$$\begin{aligned}
\frac{\partial}{\partial z}\psi_{\mathbf{k}_i}(y,z)|_{z=0} &= ik_o\, e^{i\kappa_o y} - \int d\kappa\; ik_z(\kappa)\, e^{i\kappa y}\, A_{\mathbf{k}_i}(\kappa)\\
&= \sum_n \left\{ \frac{\partial}{\partial z}\xi_{nE}(y,z)|_{z=0}\Theta_{n\mathbf{k}_i} + \frac{\partial}{\partial z}\bar{\xi}_{nE}(y,z)|_{z=0}\Delta_{n\mathbf{k}_i} \right\}, \qquad (2.7)
\end{aligned}$$

and

$$\begin{aligned}
\frac{\partial}{\partial z}\psi_{\mathbf{k}_i}(y,z)|_{z=d} &= \sum_n \left\{ \frac{\partial}{\partial z}\xi_{nE}(y,z)|_{z=d}\Theta_{n\mathbf{k}_i} + \frac{\partial}{\partial z}\bar{\xi}_{nE}(y,z)|_{z=d}\Delta_{n\mathbf{k}_i} \right\},\\
&= \int d\kappa\; ik_z(\kappa)\, e^{ik_z(\kappa)d} e^{i\kappa y}\, B_{\mathbf{k}_i}(\kappa). \qquad (2.8)
\end{aligned}$$

In order to find the wave function $\psi_{\mathbf{k}_i}(y,z)$, the coefficients $A_{\mathbf{k}_i}(\kappa)$, $B_{\mathbf{k}_i}(\kappa)$, $\Theta_{n\mathbf{k}_i}$ and $\Delta_{n\mathbf{k}_i}$ have to be determined. Equations 2.5-2.8 form a set of linear equations

satisfied by these coefficients, and in principle they can be solved to find $\psi_{\mathbf{k}_i}$. At this point, we proceed by taking the transverse Fourier transform of these equations in order to avoid the explicit y dependence in the final expressions. The transverse Fourier transform of a function $f(y)$ is defined via the equation

$$F(q) = (2\pi)^{-1/2} \int_{-\infty}^{\infty} dy\ e^{-iqy}\ f(y). \tag{2.9}$$

The transverse Fourier transforms of Equations 2.5 and 2.6 give

$$(2\pi)^{1/2}\ [\delta(q-\kappa_o) + A_{\mathbf{k}_i}(q)] = \sum_n \{\Xi_{nE}(q,0)\Theta_{n\mathbf{k}_i} + \bar{\Xi}_{nE}(q,0)\Delta_{n\mathbf{k}_i}\}, \tag{2.10}$$

$$(2\pi)^{1/2}\ e^{ik_z(q)d}\ B_{\mathbf{k}_i}(q) = \sum_n \{\Xi_{nE}(q,d)\Theta_{n\mathbf{k}_i} + \bar{\Xi}_{nE}(q,d)\Delta_{n\mathbf{k}_i}\}, \tag{2.11}$$

and those of Equations 2.7 and 2.8 give

$$(2\pi)^{1/2}\ [ik_o\ \delta(q-\kappa_o) - ik_z(q)\ A_{\mathbf{k}_i}(q)] = \sum_n \{\Xi'_{nE}(q,0)\Theta_{n\mathbf{k}_i} + \bar{\Xi}'_{nE}(q,0)\Delta_{n\mathbf{k}_i}\}, \tag{2.12}$$

$$(2\pi)^{1/2}\ e^{ik_z(q)d}\ ik_z(q)\ B_{\mathbf{k}_i}(q) = \sum_n \{\Xi'_{nE}(q,d)\Theta_{n\mathbf{k}_i} + \bar{\Xi}'_{nE}(q,d)\Delta_{n\mathbf{k}_i}\}, \tag{2.13}$$

where Ξ and Ξ' are used to denote the transverse Fourier transform of the subband wave function and its derivative along z-direction, respectively. Finally, the coefficients of plane waves in the 2D EG, $A_{\mathbf{k}_i}$ and $B_{\mathbf{k}_i}$, can be eliminated to yield

$$\begin{aligned}(2\pi)^{1/2}\ 2k_o\ \delta(q-\kappa_o) = \sum_n \{&[k_z(q)\Xi_{nE}(q,0) - i\Xi'_{nE}(q,0)]\Theta_{n\mathbf{k}_i} \\ &+[k_z(q)\bar{\Xi}_{nE}(q,0) - i\bar{\Xi}'_{nE}(q,0)]\Delta_{n\mathbf{k}_i}\},\end{aligned} \tag{2.14}$$

and

$$\sum_n \{[k_z(q)\ \Xi_{nE}(q,d) + i\Xi'_{nE}(q,d)]\Theta_{n\mathbf{k}_i} + [k_z(q)\ \bar{\Xi}_{nE}(q,d) + i\bar{\Xi}'_{nE}(q,d)]\Delta_{n\mathbf{k}_i}\} = 0. \tag{2.15}$$

Equations 2.14 and 2.15 have to be solved to obtain the coefficients Θ_n and Δ_n for a given incident plane wave with wave vector $\mathbf{k}_i$. Note that Equation 2.14

stands for the transmission of the incident plane wave into the subband states at the entrance of the constriction ($z = 0$), and Equation 2.15 for the reflection of the subband states at the exit of the constriction ($z = d$). That is, after solving the Schrödinger equation, Equation 2.3, for the subband wave functions the problem reduces to calculation of the multiple reflections from the edges of the constriction, which can be done using algebraical methods. In Sections 2.2 and 2.3 the solution procedure for these equations is described in detail.

We next determine the total current passing through the constriction assuming that the wave function $\psi_{\mathbf{k}_i}$ is calculated throughout the system. It is clear that calculating the current I_{z_o} passing through the line $z = z_o$ parallel to the z-axis one finds the current passing through system, that is through the electrical circuit connected to the QPC. Especially, one can choose z_o to lie in the constriction ($0 < z_o < d$), so that the final current expression consists of the coefficients of the subband wave functions. This way the quasi-1D nature of the system is incorporated into the current expression, and the current passing through the QPC is related to the occupation of the subbands, as the Landauer formulae and the above simple explanation in Section 1.2 conjecture. The current due to the incident waves with energy E can be written using the expectation value of the current operator

$$< f|\hat{j}|g >= \frac{\hbar}{2im^*} \int dy \ [< f|\partial g/\partial z > - < \partial f/\partial z|g >] \tag{2.16}$$

with respect to the current carrying solutions $\psi_{\mathbf{k}_i}(y, z)$ for all of the incident plane waves

$$J(E) = 2e \int \frac{d\mathbf{k}_i}{(2\pi)^2} \ < \psi_{\mathbf{k}_i}(y, z)|\hat{j}|\psi_{\mathbf{k}_i}(y, z) > |_{z=z_o} \ \delta \left(\frac{\hbar^2 |\mathbf{k}_i|^2}{2m^*} - E \right) \ \theta(k_o). \tag{2.17}$$

where the factor e is included to convert the probability current into the electrical current, the prefactor 2 takes care of the spin degeneracy and $1/(2\pi)^2$ is the density of states in the 2D k-space. The δ-function selects the states with energy E and $\theta(k_o)$ guarantees that the incident wave vector $\mathbf{k}$ is pointing towards the constriction. Using the wave function given in Equation 2.4 one can express the

current as

$$J(E) = \frac{e}{\pi h}\int_{-k_E}^{k_E}\frac{d\kappa_o}{k_o}\,\mathrm{Im}\sum_{nm}\;\Big\{\Theta^*_{n\mathbf{k}_i}[\int dy\xi^*_{nE}(y,z_o)\frac{\partial}{\partial z}\xi_{mE}(y,z)|_{z=z_o}]\Theta_{m\mathbf{k}_i}$$
$$+\Delta^*_{n\mathbf{k}_i}[\int dy\bar{\xi}^*_{nE}(y,z_o)\frac{\partial}{\partial z}\bar{\xi}_{mE}(y,z)|_{z=z_o}]\Delta_{m\mathbf{k}_i}$$
$$+\Theta^*_{n\mathbf{k}_i}[\int dy\xi^*_{nE}(y,z_o)\frac{\partial}{\partial z}\bar{\xi}_{mE}(y,z)|_{z=z_o}]\Delta_{m\mathbf{k}_i}$$
$$+\Delta^*_{n\mathbf{k}_i}[\int dy\bar{\xi}^*_{nE}(y,z_o)\frac{\partial}{\partial z}\xi_{mE}(y,z)|_{z=z_o}]\Theta_{m\mathbf{k}_i}\Big\}, \tag{2.18}$$

with $E = \hbar^2 k_E/2m^*$. The total current passing through the constriction is calculated by integrating the current energy density $J(E)$ over the whole energy range, using a weighting factor which is the number of electrons moving to right that are in excess of those moving to left. We assume that two reservoirs are connected to the 2D EG at $z = \pm\infty$ and the electrochemical potentials of these reservoirs are kept constant so that there is an infinitesimal difference $\Delta\mu = \mu_L - \mu_R$ between the electrochemical potential of the left- (μ_L) and right-hand side (μ_R) reservoirs. Thus, the weighting factor becomes $[f^L_{FD}(E) - f^R_{FD}(E)]$, f_{FD} denoting the Fermi-Dirac distribution function and the superscript indicating the reservoir for which it is calculated. The conductance of the QPC, on the other hand, depends on how the voltage in the circuit is measured as discussed in Section 1.1. We are adopting a two-terminal measurement geometry. That is, the voltage measured by the electrical circuit connected to the QPC is just the electrochemical potential difference between the reservoirs $V = \Delta\mu/e$. Then, the conductance is defined as the ratio of the current passing through the constriction to the difference of voltages measured deep in the reservoirs. At zero temperature, only the states lying at the Fermi level contribute the current, and for infinitesimal bias one has

$$I = \int dE\;[f_{FD}(E) - f_{FD}(E+\Delta\mu)]\;J(E) = J(E_F)\;\Delta\mu = J(E_F)\;eV. \tag{2.19}$$

Finally, the conductance is given by

$$G = \frac{I}{V} = e\;J(E_F). \tag{2.20}$$

That is, the conductance of the QPC in a two-terminal geometry is given by Equation 2.18 multiplied by e. Since in the experiments the conductance is measured in a two-terminal configuration, the effects of the self-consistent field of the nonequilibrium electrons, and the resulting electrochemical potential difference between the two sides of the constriction is not reflected in the results. The treatment given above is consistent with the experimental results as far as the relevant Landauer formula is concerned. Earlier, Landauer[77] conjectured that the dilution at the wide regions connected to the QPC have the effect of a reservoir except the phase randomization. Numerical calculations by Yosefin and Kaveh[80] showed that this effect is present for constrictions with smoothly varying cross-section. The model used in this study, however, has an abrupt and infinite jump in cross-section at the edges of the constriction. Therefore, the dilution effect is even stronger for this case and making four-terminal measurements is possible only by including the voltage probes in the constriction, i.e., using a cross-geometry. For the electron waveguide geometry used for observing the quantization of conductance, however, the measurement is a two-terminal one.

The experimental observation of quantization of conductance[58,59] in QPC lead to numerous theoretical studies[81–92] to appear on the subject. In these studies various approaches and techniques have been used, all leading to qualitatively similar results. This can be taken as a justification of the generic nature of the quantization of conductance and its independence of the details of the system. The techniques used in these studies were widespread and verifying that the quantization is robust against several assumptions exploited to investigate the system theoretically. In addition, the robustness of quantization against the effects of constriction geometry,[84,88,93,94] temperature averaging,[84,88,95,96] adiabatic evolution of states,[81,88,89] channel roughness,[94] impurity scattering[97–100] and disorder in the channel[87,101] was studied; several related systems such as QPC in series[102] and parallel,[103] quasi-0D electron cavities[94] and 1D crystals[104] were investigated; and the relation between the Landauer formulae and QPC systems[80,105,106] was analyzed. We wish to summarize the differences and similarities among the methods used in these studies. These methods can

be classified into two main groups, namely as those using the quasi-one-dimensionality of the system in an explicit way[82,84,88] and those aiming at solving the full-Schrödinger equation,[80,81,85–87,89] Equation 2.3. Those in the first group result in almost the same formulation of the problem, which at the end yield Equation 2.18. The second group consists of the exact[80,81,92] and approximate[89] solutions of Equation 2.3 and applications of the tight-binding method,[85] scattering theoretical methods[86] and Anderson model[87,101] to the problem of the QPC. The first approach is shown[82,88,94,95,107] to provide simplicity and versatility and thus it is more appropriate to carry out a systematical study of the quantization of conductance. In the following sections and chapters the wide range of applications of the present formalism is exemplified in detail.

2.2 Uniform Constriction

2.2.1 Formalism

The existing studies pointed out the fact that approximating the QPC as a uniform channel between two 2D EG with abrupt junctions is not a realistic approach. Nevertheless, the uniform constriction is the easiest one to solve and it still has some ingredients of the real system.[88] In this section the general formalism described above is exemplified by use of the uniform constriction.

Here the word uniform refers to the fact that the confining potential is the same throughout the constriction and the longitudinal part of the potential vanishes. Consequently, the Schrödinger equation, Equation 2.3, takes the form

$$\left[\left\{-\frac{\hbar^2}{2m^*}\frac{\partial^2}{\partial z^2}\right\}+\left\{-\frac{\hbar^2}{2m^*}\frac{\partial^2}{\partial y^2}+V_c(y)\right\}\right]\xi_{nE}(y,z)=E\xi_{nE}(y,z), \qquad (2.21)$$

for the uniform constriction. This equation is separable and its solutions are given by

$$\xi_{nE}(y,z)=e^{i\gamma_n z}\,\phi_n(y), \qquad (2.22)$$

where the lateral wave function $\phi_n(y)$ is defined by the equation

$$\left[-\frac{\hbar^2}{2m^*}\frac{d^2}{dy^2} + V_c(y)\right]\phi_n(y) = \epsilon_n\phi_n(y), \tag{2.23}$$

with subband energy ϵ_n. The propagation constant along the z-direction is given by

$$\gamma_n = \sqrt{\frac{2m^*}{\hbar^2}(E - \epsilon_n)}, \tag{2.24}$$

where the root with the positive imaginary part is chosen. The lateral Fourier transforms Ξ and Ξ' used in Equations 2.5-2.8 can easily be found in terms of the lateral Fourier transform of the subband wave function $\Phi_n(q)$. In turn, Equations 2.14 and 2.15 can be written as

$$(2\pi)^{1/2}\ 2k_o\ \delta(q-\kappa_o) = \sum_n\{[k_z(q)+\gamma_n]\Theta_{n\mathbf{k}_i} + [k_z(q)-\gamma_n]\Delta_{n\mathbf{k}_i}\}\Phi_n(q), \tag{2.25}$$

and

$$0 = \sum_n\{[k_z(q)-\gamma_n]\ e^{i\gamma_n d}\Theta_{n\mathbf{k}_i} + [k_z(q)+\gamma_n]\ e^{-i\gamma_n d}\Delta_{n\mathbf{k}_i}\}\Phi_n(q). \tag{2.26}$$

Unlike Equations 2.14 and 2.15, these equations can easily be solved. Multiplying these equations from left by $\Phi_m^*(q)$ and integrating over all q one gets

$$(2\pi)^{1/2}\ 2k_o\ \Phi_m^*(\kappa_o) = \sum_n\{[K_{mn}+\delta_{mn}\gamma_n]\Theta_{n\mathbf{k}_i} + [K_{mn}-\delta_{mn}\gamma_n]\Delta_{n\mathbf{k}_i}\}, \tag{2.27}$$

and

$$0 = \sum_n\{[K_{mn}-\delta_{mn}\gamma_n]\ e^{i\gamma_n d}\Theta_{n\mathbf{k}_i} + [K_{mn}+\delta_{mn}\gamma_n]\ e^{-i\gamma_n d}\Delta_{n\mathbf{k}_i}\}, \tag{2.28}$$

where δ_{mn} is the Kronecker delta function and K_{mn} is defined via

$$K_{mn} = \int dq\ \Phi_m^*(q)\ k_z(q)\ \Phi_n(q). \tag{2.29}$$

Deriving these relations we make use of the orthonormality of the lateral wave functions $\phi_n(y)$ and thus their Fourier transforms

$$\int dq\ \Phi_m^*(q)\ \Phi_n(q) = \delta_{mn}. \tag{2.30}$$

Clearly, Equations 2.27 and 2.28 can be cast in a matrix form as

$$(2\pi)^{1/2}\ 2k_o\ \tilde{\Phi}^{\dagger}(\kappa_o) = \{[\tilde{K} + \tilde{\Gamma}]\tilde{\Theta}_{\mathbf{k}_i} + [\tilde{K} - \tilde{\Gamma}]\tilde{\Delta}_{\mathbf{k}_i}\}, \tag{2.31}$$

and

$$0 = \{[\tilde{K} - \tilde{\Gamma}]\ e^{i\tilde{\Gamma}d}\tilde{\Theta}_{\mathbf{k}_i} + [\tilde{K} + \tilde{\Gamma}]\ e^{-i\tilde{\Gamma}d}\tilde{\Delta}_{\mathbf{k}_i}\}. \tag{2.32}$$

Here $\tilde{\Theta}$ and $\tilde{\Delta}$ are the column vectors of the coefficients of the subband wave functions with right- and left-going probability currents, respectively. $\tilde{\Phi}$ is the row vector of the transverse Fourier transform of the lateral wave functions. $\tilde{\Gamma}$ is the matrix of propagation constants and $\tilde{K}$ is the longitudinal momentum matrix with elements $\Gamma_{mn} = \delta_{mn}\gamma_n$ and K_{mn}, respectively.

Before obtaining the solution for the coupled linear equations Equations 2.31 and 2.32 we wish to focus our attention on the physical interpretation of these equation. Considering only the junction at $z = 0$ (i.e., between the left-hand side 2D EG and the constriction) and assuming that there are no left-going states occupied in the constriction, Equation 2.31 becomes

$$(2\pi)^{1/2}\ 2k_o\ \tilde{\Phi}^{\dagger}(\kappa_o) = [\tilde{K} + \tilde{\Gamma}]\tilde{t}_{\mathbf{k}_i}, \tag{2.33}$$

with the solution

$$\tilde{t}_{\mathbf{k}_i} = (2\pi)^{1/2}\ 2k_o\ [\tilde{K} + \tilde{\Gamma}]^{-1}\ \tilde{\Phi}^{\dagger}(\kappa_o). \tag{2.34}$$

Here we used $\tilde{t}$ instead of $\tilde{\Theta}$ since the problem at hand corresponds to transmission into the constriction for an incident plane wave from the 2D EG, and the coefficients of the subband wave functions are just the corresponding transmission probability amplitudes. The vector $\tilde{t}$ is analogous to the transmission amplitude for a 1D infinite barrier $t = 2k/(k + k')$. The only difference is that the presence of the subband structure has to be taken into account and all the relevant quantities in the strictly 1D problem has to be converted into matrices by using $\{\phi_m(y)\}$ as a basis. Next, considering only the junction at $z = d$ and assuming that the occupation of the right-going states in the constriction is given by an arbitrary vector $\tilde{x}$, Equation 2.32 becomes

$$0 = \{[\tilde{K} - \tilde{\Gamma}]\ \tilde{x} + [\tilde{K} + \tilde{\Gamma}]\ \tilde{r}\tilde{x}\}, \tag{2.35}$$

where the square matrix $\tilde{\mathbf{r}}$ includes the reflection probability amplitudes back into the constriction for incident subband wave functions and is given by

$$\tilde{\mathbf{r}} = [\tilde{\mathbf{K}} + \tilde{\mathbf{\Gamma}}]^{-1} \, [\tilde{\mathbf{\Gamma}} - \tilde{\mathbf{K}}]. \tag{2.36}$$

Similar to the transmission amplitude, the reflection amplitude is analogous to the strictly 1D equivalent $r = (k - k')/(k + k')$.

Clearly, this analogy with the 1D case holds for all of the subsequent formulation. For example, the wave function for a finite length constriction can be visualized by using the result for a finite rectangular barrier in 1D. The amplitude of the right-going wave in the barrier for the 1D problem is given by $\vartheta = t/[1 - r^2 \exp(2ik'd)]$ with the above t and r. Apparently, the analogy does not allows one to write the solution without making any calculations since the matrix multiplication is not commutative. Nevertheless, it is beneficial when interpreting the results. Moreover, a simple 1D picture simplifies the underlying physics to a great extent. Returning back to the problem of uniform finite constriction, Equations 2.31 and 2.32 can be solved simultaneously to yield

$$\tilde{\mathbf{\Delta}}_{\mathbf{k}_i} = e^{i\tilde{\mathbf{\Gamma}}d} \, \tilde{\mathbf{r}} \, e^{i\tilde{\mathbf{\Gamma}}d} \, \tilde{\mathbf{\Theta}}_{\mathbf{k}_i}, \tag{2.37}$$

$$\tilde{\mathbf{\Theta}}_{\mathbf{k}_i} = [\tilde{\mathbf{1}} - \tilde{\mathbf{r}} e^{i\tilde{\mathbf{\Gamma}}d} \tilde{\mathbf{r}} e^{i\tilde{\mathbf{\Gamma}}d}]^{-1} \, \tilde{\mathbf{t}}_{\mathbf{k}_i}. \tag{2.38}$$

This completes the determination of the wave function in the constriction. Next we calculate the conductance of the uniform constriction. Using the fact that $\{\phi_m(y)\}$ form an orthonormal complete set, Equation 2.18 reduces to the simple expression

$$\begin{aligned} J(E) = \frac{e}{\pi h} \int_{-k_E}^{k_E} \frac{d\kappa_o}{k_o} \, \mathrm{Im} \, \Big\{ & \tilde{\mathbf{\Theta}}^{\dagger}_{\mathbf{k}_i} \, e^{-i\tilde{\mathbf{\Gamma}}^{\dagger} z_o} i\tilde{\mathbf{\Gamma}} e^{i\tilde{\mathbf{\Gamma}} z_o} \tilde{\mathbf{\Theta}}_{\mathbf{k}_i} - \tilde{\mathbf{\Delta}}^{\dagger}_{\mathbf{k}_i} e^{i\tilde{\mathbf{\Gamma}}^{\dagger} z_o} i\tilde{\mathbf{\Gamma}} e^{-i\tilde{\mathbf{\Gamma}} z_o} \tilde{\mathbf{\Delta}}_{\mathbf{k}_i} \\ & - \tilde{\mathbf{\Theta}}^{\dagger}_{\mathbf{k}_i} e^{i\tilde{\mathbf{\Gamma}}^{\dagger} z_o} i\tilde{\mathbf{\Gamma}} e^{-i\tilde{\mathbf{\Gamma}} z_o} \tilde{\mathbf{\Delta}}_{\mathbf{k}_i} + \tilde{\mathbf{\Delta}}^{\dagger}_{\mathbf{k}_i} e^{i\tilde{\mathbf{\Gamma}}^{\dagger} z_o} i\tilde{\mathbf{\Gamma}} e^{i\tilde{\mathbf{\Gamma}} z_o} \tilde{\mathbf{\Theta}}_{\mathbf{k}_i} \Big\}. \end{aligned} \tag{2.39}$$

The imaginary part of the expression in the brackets can easily be calculated and comes out to be independent of z_o as pointed out above. The final form of the conductance is given by

$$G = \frac{e^2}{\pi h} \int_{-k_F}^{k_F} \frac{d\kappa_o}{k_o} \{ \tilde{\mathbf{\Theta}}^{\dagger}_{\mathbf{k}_i} \tilde{\mathbf{\Gamma}}_R \tilde{\mathbf{\Theta}}_{\mathbf{k}_i} - \tilde{\mathbf{\Delta}}^{\dagger}_{\mathbf{k}_i} \tilde{\mathbf{\Gamma}}_R \tilde{\mathbf{\Delta}}_{\mathbf{k}_i} + 2\mathrm{Im}[\tilde{\mathbf{\Theta}}^{\dagger}_{\mathbf{k}_i} \tilde{\mathbf{\Gamma}}_I \tilde{\mathbf{\Delta}}_{\mathbf{k}_i}] \}, \tag{2.40}$$

where $\tilde{\Gamma}_R = \text{Re}\,\{\tilde{\Gamma}\}$ and $\tilde{\Gamma}_I = \text{Im}\,\{\tilde{\Gamma}\}$ so that $\tilde{\Gamma} = \tilde{\Gamma}_R + i\tilde{\Gamma}_I$. Here three terms in the brackets have different characters. The first and second terms correspond to right- and left-going wave in the constriction, respectively. They may lead to some resonances as a result of the differing relative phase as shown later in this section. The third term, on the other hand, corresponds to evanescent states in the constriction. For finite constrictions this term is of importance and this term yields an important deviation from the exact quantization. As pointed out before, the conductance of the QPC is expressed in terms of the occupation of the subbands. Nevertheless, Equation 2.40 can not be thought as a Landauer formula, since the cross-section for the system under consideration is changing discontinuously. Therefore, neither the results for a uniform system nor those obtained by an adiabatic approximation are applicable. However, to be precise, we have to stress that the measurement geometry corresponds to the Landauer formula $G \sim T$.

2.2.2 Results

The conductance expression given in Equation 2.40 can not be compared directly with the experimental results.[58,59] In the experiments the gate voltage V_g is changed and the corresponding conductance $G(V_g)$ is measured. Equation 2.40, however, gives the conductance as a function of electron density (or equivalently λ_F) and the width of the constriction w. Therefore one has to calculate λ_F and w for a given split-gate geometry and V_g in order to have a direct comparison. This is done both self-consistently[108] and empirically[109] for infinite channels. According to the self-consistent solutions,[108] when the number of the occupied subbands is small $V_c(y)$ is parabolic, and as the number increases it becomes flat near the center in the y-direction. This result was verified for the QPC used in the experiments,[109] and it was shown that the most appropriate form for $V_c(y)$ is a quantum well with parabolic walls. Note that, the essential features responsible for the quantization of conductance are independent of the *form of the confining potential.* Therefore, we focus our attention only on these generic results and do

not consider the complications arising from the details of the electrostatic field and self-consistent field. Later we devise an approximate scheme to investigate the possible effects originating therefrom.

In this section we apply the present formalism to both infinite-well confinement

$$\begin{aligned} V_c(y) &= 0 \qquad |y| < w/2, \\ &= \infty \qquad \text{otherwise}, \end{aligned} \tag{2.41}$$

and parabolic confinement

$$V_c(y) = \frac{1}{2} m^* \omega^2 y^2, \tag{2.42}$$

since they provide important simplifications in calculations and resemble to the realistic solution to some extent. The subband energies for these confining potentials are given by

$$\begin{aligned} \epsilon_n &= \frac{\hbar^2}{2m^*} \left(\frac{n\pi}{w}\right)^2 \qquad \text{infinite} - \text{well}, \\ &= \hbar\omega \left(n + \frac{1}{2}\right) \qquad \text{parabolic}. \end{aligned} \tag{2.43}$$

The corresponding lateral wave functions are sine waves and Hermite polynomials multiplied by Gaussians, respectively. Their Fourier transforms are again well known and we give the results only. For the infinite-well confinement

$$\Phi_n(q) = i^{n+1} \sqrt{2w} f_n\left(\frac{qw}{2}\right) \frac{n}{(qw)^2 - (n\pi)^2}, \tag{2.44}$$

where $f_n(x) = \cos x$ for $n =$ odd and $f_n(x) = -\sin x$ for $n =$ even, and for the parabolic confinement

$$\Phi_n(q) = \frac{(-i)^n}{\sqrt{2^n \pi^{1/2} n!}} \, w^{1/2} \, H_n(qw) e^{-(qw)^2/2}, \tag{2.45}$$

where $H_n(x)$ is the Hermite polynomial of degree n and the natural length of the harmonic oscillator is $w = \sqrt{\hbar/m^*\omega}$. Note that, since the confining potential is a symmetric function of y, the lateral wave functions and consequently their Fourier transforms are either even or odd functions. That is, $\Phi_n(q)$ is either real or

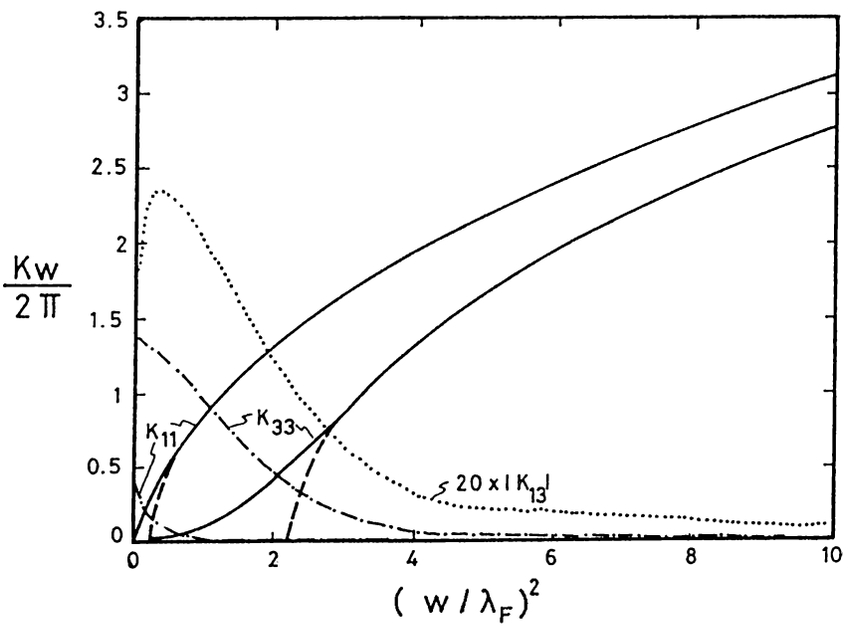

Figure 2.1: Longitudinal momentum matrix elements for infinite-well confinement
The full and dash-dotted lines correspond to the real and imaginary parts, respectively, of K_{11} and K_{33} as a function of width. The dashed lines show the real part of the propagation constants γ_1 and γ_3. The dotted line is the magnitude of K_{13} amplified 20 times.

imaginary. The longitudinal momentum matrix $\tilde{\mathbf{K}}$ can be calculated by numerical integration by using these functions. Clearly the matrix element for two wave functions with different symmetries is zero. That is, the even and odd index subbands are completely decoupled. In Figures 2.1 and 2.2 some elements of $\tilde{\mathbf{K}}$ are shown as a function of width w for infinite-well and parabolic confinements, respectively. An important observation is that the off-diagonal elements of $\tilde{\mathbf{K}}$ are very small compared to the diagonal elements, especially for the infinite-well confinement. The off-diagonal elements become appreciable only when the energy of a subband dips below the Fermi level. For comparison the propagation constants γ_n are also shown in the figures. For large w values the diagonal longitudinal matrix element K_{nn} approaches to the propagation constant γ_n. On

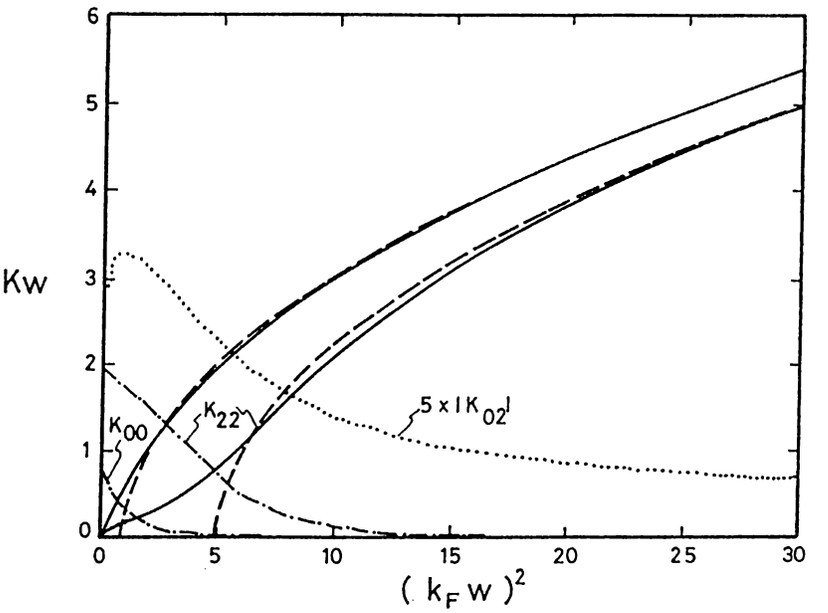

Figure 2.2: Longitudinal momentum matrix elements for parabolic confinement The full and dash-dotted lines correspond to the real and imaginary parts, respectively, of K_{00} and K_{22} as a function of width. The dashed lines show the real part of the propagation constants γ_0 and γ_2. The dotted line is the magnitude of K_{02} amplified 5 times.

the other hand, for smaller w they differ as result a of quantum size effect, i.e., the momentum spectrum of the lateral wave function gets wider and the plane waves with energy E become insufficient to compose this wave function.

Next, we examine the transmission into and out of a semiinfinite constriction. For incidence from the 2D EG the total current going into the constriction is found by eliminating the terms due to the left-going and evanescent states from Equation 2.40

$$G_t = \frac{e^2}{\pi h} \int_{-k_F}^{k_F} \frac{d\kappa_o}{k_o} \, \tilde{t}^{\dagger}_{k_i} \tilde{\Gamma}_R \tilde{t}_{k_i}, \tag{2.46}$$

and using Equation 2.34 one gets

$$G_t = \frac{2e^2}{h} \int_{-k_F}^{k_F} d\kappa_o \; 4k_o \tilde{\Phi}(\kappa_o) \; ([\tilde{K} + \tilde{\Gamma}]^{-1})^{\dagger} \tilde{\Gamma}_R [\tilde{K} + \tilde{\Gamma}]^{-1} \tilde{\Phi}^{\dagger}(\kappa_o). \tag{2.47}$$

In the integrand only k_o and $\tilde{\Phi}$ are functions of κ_o and by using Equation 2.29

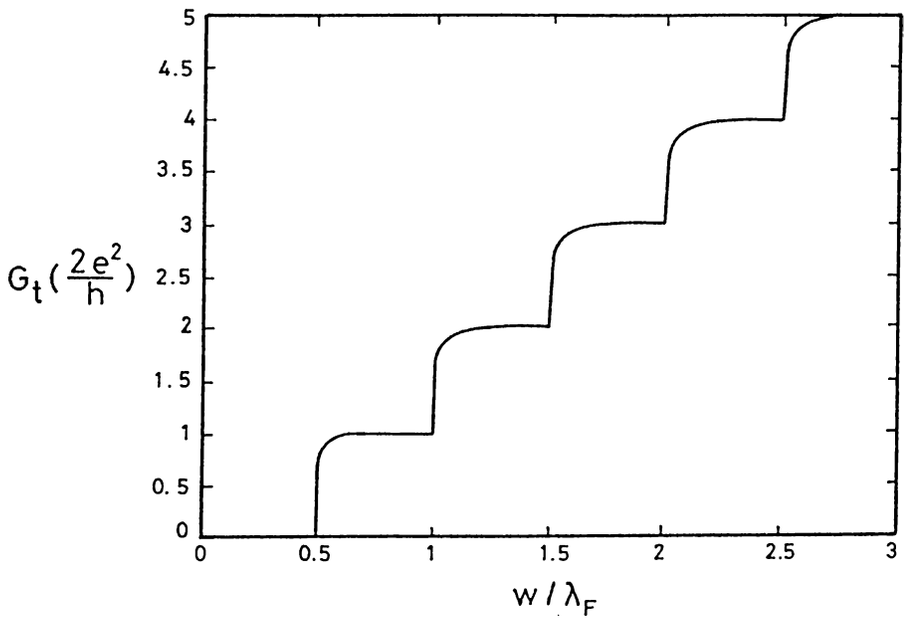

Figure 2.3: Conductance due to transmission into a semiinfinite uniform channel The conductance as a function of width. Infinite-well confinement.

one gets

$$\mathrm{Re}\,\tilde{\mathrm{K}} = \int_{-k_F}^{k_F} d\kappa\ \tilde{\Phi}^{\dagger}(\kappa)\ k_z(\kappa)\ \tilde{\Phi}(\kappa), \tag{2.48}$$

since $k_z(\kappa)$ is real for $|\kappa| < k_F$. Finally, the conductance due to transmission into a semiinfinite constriction can be written as

$$G_t = \frac{2e^2}{h}\mathrm{Tr}\left\{4([\tilde{\mathrm{K}} + \tilde{\Gamma}]^{-1})^{\dagger}\tilde{\Gamma}_R[\tilde{\mathrm{K}} + \tilde{\Gamma}]^{-1}\mathrm{Re}(\tilde{\mathrm{K}})\right\}. \tag{2.49}$$

This way the integration over the incident wave vectors is eliminated and the result is expressed only in terms of matrices[‡]. In Figure 2.3 the conductance

[‡]The important point is that $\tilde{\mathrm{K}}$ and $\tilde{\Gamma}$ are infinite dimensional matrices. For a numerical calculation, however, the matrices used have to be finite dimensional. This problem, actually, arises in the calculation of both $\tilde{\mathrm{t}}$ and $\tilde{\mathrm{r}}$ and can be overcome by using an approximation scheme. Using finite dimensional truncations of $\tilde{\mathrm{K}}$ and $\tilde{\Gamma}$ the corresponding conductance is calculated. The dimension of these matrices are increased until a reasonable convergence for the conductance is obtained. In fact this convergence is obtained fairly easily since the off-diagonal elements of $\tilde{\mathrm{K}}$ is very small compared to the diagonal elements. For example, the results presented in this chapter are obtained at most 10 subbands and it is usually sufficient to include one or two subbands above E_F.

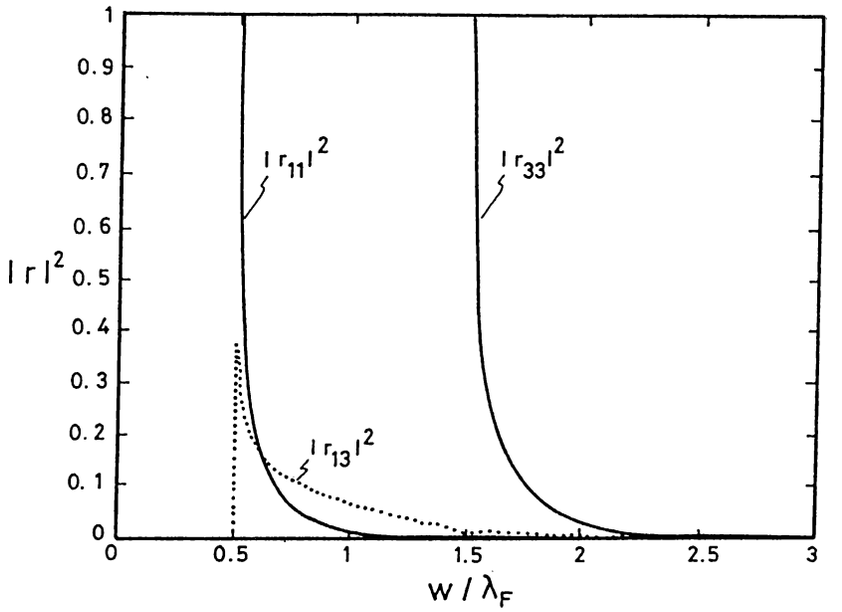

Figure 2.4: Reflection matrix elements
The squared magnitude of r_{11} and r_{33} are shown by full lines as a function of width. The dotted curve corresponds to squared magnitude of r_{13} and shown only when the first subband is occupied. Infinite-well confinement.

G_t is shown for infinite-well confinement, in which the approximate quantization of conductance is apparent. This point is understandable, since the trace in Equation 2.49 approaches to N_p as $K_{nn} \to \gamma_n$. Similarly, for incidence from the constriction the reflection amplitudes are given by Equation 2.36. The important difference between the two cases is that for the latter one may calculate the conductance just by using conventional Landauer formula in the channel

$$G_r = \frac{2e^2}{h}\left[N_p - \mathrm{Tr}\{\tilde{\mathbf{r}}_p^{\dagger}\tilde{\mathbf{r}}_p\}\right], \tag{2.50}$$

where the first term in the brackets is just the number of occupied subbands N_p in the constriction and gives the contribution of the incident waves. The second term is the contribution of the reflected waves. Clearly, $G_t = G_r$ as a result of the time-reversal symmetry. In Figure 2.4 the reflection amplitudes r_{mn} are shown for infinite well confinement.

After examining the semiinfinite constriction we now study the constriction of finite length. Using Equation 2.37 and 2.48 it is possible to write Equation 2.40 as a trace as it is done for G_t. The final expression is complicated and is not given here. The results obtained for finite length constrictions are shown in Figure 2.5 and 2.6 for infinite-well and parabolic confinements, respectively. In the following discussion the infinite-well confinement is used as the reference.

The zero-length QPC is one of the classically and semiclassically treated problems as described in Section 1.2. Comparing the result obtained by using the present formalism with those of Equations 1.24 and 1.30 one observes that the full quantum mechanical conductance G resembles the Sharvin conductance G_s more than the crude quantum mechanical result G_p. The result is not surprising, since for an infinitesimally long QPC the probability of tunneling is appreciable and for infinitely long one it is exactly zero. Nevertheless, G for $d = 0$ as a function of w shown in Figure 2.5 is not a straight line passing through the origin. It is shifted towards larger w values and weak oscillations are superimposed on it. The reason for the shift is the Heisenberg's uncertainity principle. According to $\Delta y \Delta p_y \gtrsim \hbar$ the transverse momentum has to increase as $\Delta y \simeq w$ decreases. However, the largest possible $|p_y|$ is $(\hbar k_F)$ and consequently for very narrow constrictions (i.e., $k_F w \lesssim 1$) transport is suppressed. The oscillations, on the other hand, may be thought as the precursors of quantized conductance. For $\epsilon_n > E_F$ the transport is via tunneling and conductance increases exponentially or at worse polynomially. A subband dips in the Fermi level for $\epsilon_n < E_F$ and the nature of transport changes to free propagation. The maximum conductance for a subband is limited by the quantum of conductance $(2e^2/h)$, however. Therefore, the conductance due to this single subband saturates leading to the formation of weak shoulder-like features in Figure 2.5. It is also worthwhile to visualize these oscillations as the diffraction effects of a slit in an optical system.

For finite length, the tunneling contribution decreases as an exponential function of d since as found from Equations 2.37 and 2.38 $\tilde{\Delta} \sim \exp(-2\tilde{\Gamma}_I d)\tilde{\Theta}$ and $\tilde{\Theta} \sim \tilde{t}$. Consequently the steps get sharper. The forementioned oscillations superimposed on the classical Sharvin conductance, on the other hand, evolve

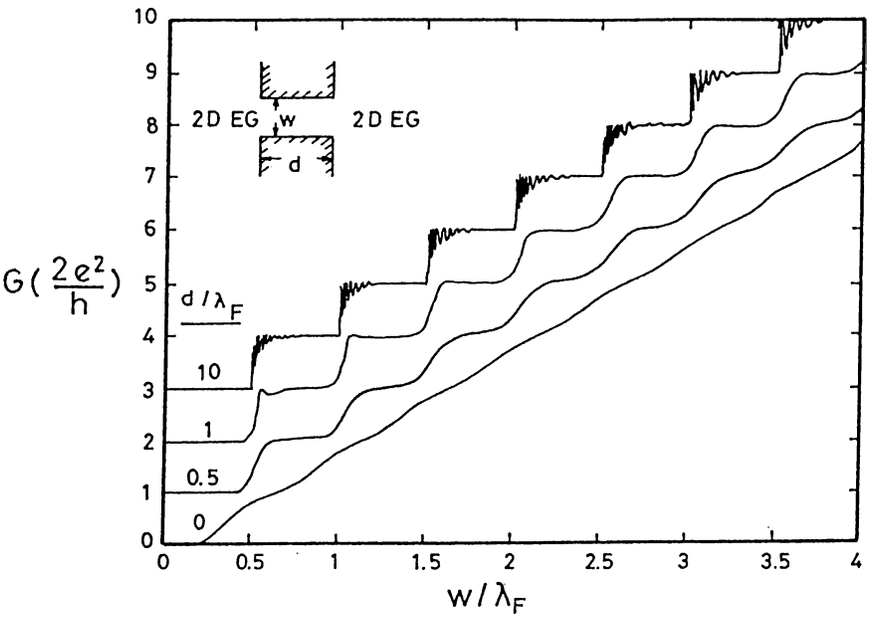

Figure 2.5: Conductance of quantum point contacts with infinite-well confinement
The conductance as a function of width for several values of constriction length, given in units of λ_F. The curves are offset for clarity. The inset shows the geometry used in the calculations.

to form the quantized plateaus for $d \gtrsim \lambda_F/2$. The quantization gets better with increasing d and for $d \gtrsim 5\lambda_F$ the plateaus occur exactly at multiples of $2e^2/h$ within the precision of computations. However, the form of the conductance curve is still quite different than that conjectured by Equation 1.30 and that observed in the experiments.[58,59] Namely, the conductance does not form flat plateaus but display oscillations below the quantized values. These are the resonances caused by the interference of right- and left-going waves in the constriction. To analyze these resonances we examine Equation 2.37. The matrix $\exp(i\tilde{\Gamma}d)$ consists of pure phases for occupied subbands and varying w (i.e., varying $\tilde{\Gamma}$) these phases change as well. This yields an interference between the first and second terms in the brackets in Equation 2.40. This can be seen by approximating $\tilde{\mathrm{K}}$ by a diagonal matrix, i.e., $K_{mn} = K_n \delta_{mn}$. Then, the reflection matrix is also diagonal

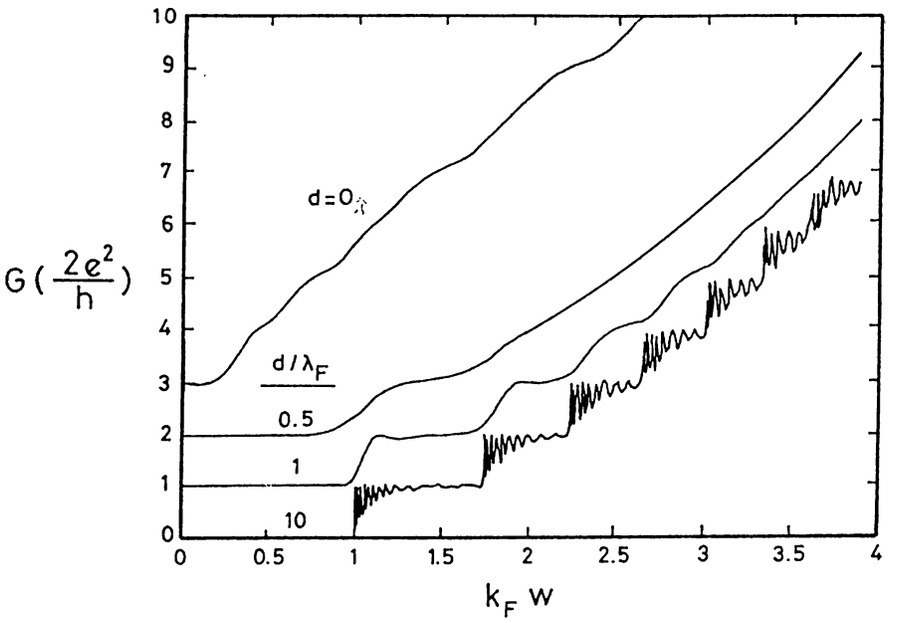

Figure 2.6: Conductance of quantum point contacts with parabolic confinement The conductance as a function of width for several values of constriction length, given in units of λ_F. The curves are offset for clarity.

$r_{mn} = r_n \delta_{mn}$ with

$$r_n = \frac{K_n - \gamma_n}{K_n + \gamma_n}, \tag{2.51}$$

and the transmission vector is

$$t_{nk_i} = (2\pi)^{1/2} \frac{2k_o}{K_n + \gamma_n} \Phi_n^*(\kappa_o). \tag{2.52}$$

Using this diagonal approximation in Equations 2.37-2.40 one obtains $G = 2e^2/h \times \sum_n T_n$ where

$$T_n = \frac{(|\gamma_n + K_n|^2 - |\gamma_n - K_n|^2)^2}{||\gamma_n + K_n|^2 - |\gamma_n - K_n|^2 e^{2i(\gamma_n d + \varphi_n)}|^2}, \tag{2.53}$$

for subbands below the Fermi level. A similar expression can be found for subbands above E_F giving the tunneling contribution. Here φ_n is the phase of r_n. This expression becomes unity for $\gamma_n d + \varphi_n = \ell\pi$ and has minima for the same quantity being equal to $(\ell + 1/2)\pi$. Note that, the envelope of the

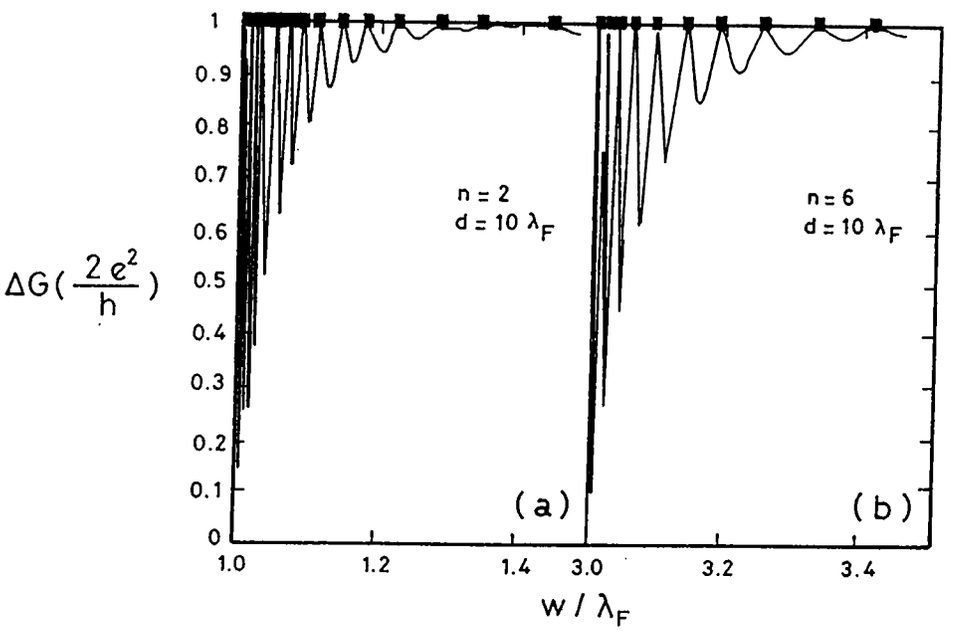

Figure 2.7: Detail of resonance structure for quantum point contacts
The conductance due to the uppermost subband below the Fermi level as a function of width. The length of the constriction is $10\times\lambda_F$. The filled squares are the resonance positions calculated from Equation 2.54 using $c = 0.965$ and 0.945, for $n = 2$ and 6, respectively. Infinite-well confinement.

antiresonances, i.e., the minima of the conductance, is independent of d for a given subband index n. A simple approximation for the position of resonances may be obtained by taking $\varphi_n = 0$, i.e., assuming K_n is real. Then, the resonance condition becomes $\gamma_n d = \ell\pi$. Using this equality the position of ℓth resonance on the nth plateau

$$w_{n\ell} \simeq \frac{n\lambda_F}{2} \frac{1}{\sqrt{1-\left(\frac{\ell\lambda_F}{2d}\right)^2}}, \tag{2.54}$$

and the number of resonances on the nth plateau

$$M_n \simeq \frac{2d}{\lambda_F} \frac{\sqrt{2n+1}}{n+1}, \tag{2.55}$$

can be calculated. Some immediate results of Equations 2.54 and 2.55 may be found such as the conductance does not reach to quantized values for $d < \lambda_F/\sqrt{3}$

and the resonances become denser as d increases or n decreases. However, taking the phase of the reflection coefficient ϕ_n equal to zero is a good approximation only just above the propagation threshold $w = n\lambda_F/2$. The results given by Equation 2.54 deviates from the exact resonance positions and even match with the antiresonance positions due to the nonzero phase. In order to prevent this deviation we use $\bar{\ell} = c\ell$ instead of ℓ in Equation 2.54, where c is a constant. These results are compared with the full calculation in Figure 2.7.

2.2.3 Effects of Finite Temperature and Bias

As mentioned above, the experimental results[58,59] do not look like what we have obtained in Figures 2.5 and 2.6. The most important discrepancy is the lack of observation of the resonance structure in the experiments. Although some fine structure on top of the quantized plateaus were observed[110] they are not regular as the above described resonances. It is known that finite temperature and bias yields to an averaging over the extended Fermi circle of the 2D EG and therefore are candidates for possible effects that may wash out the resonance structure. On the other hand, we have to investigate whether the quantization of conductance, as well, disappears with finite temperature and bias. It is concluded from the experiments[58,59] that the resonance structure has to be washed out while the steps remain for realistic systems.

In reality the effect of temperature on the quantization of conductance is twofold. First, the mobility of the sample decreases with temperature.[111] Therefore, the mean free path of electrons get shorter and the transport becomes diffusive instead of ballistic. However, for temperatures at which the experiments are carried out the temperature is still low enough that the transport is ballistic. As a first approximation one can ignore the effects of the temperature on mobility and focus on the averaging near the Fermi circle. For a finite temperature T, the total current passing through the constriction is calculated by using the first equality in Equation 2.19

$$I = \int dE \, [f_{FD}(E;T) - f_{FD}(E + \Delta\mu;T)] \, J(E). \tag{2.56}$$

Assuming that the difference of the electrochemical potential $\Delta\mu$ is small enough one can use the Taylor expansion for the term in the brackets

$$I = \Delta\mu \int dE \left(-\frac{\partial f_{FD}(E;T)}{\partial E}\right) J(E), \tag{2.57}$$

which leads to conductance

$$G_T = \int dE \left(-\frac{\partial f_{FD}(E;T)}{\partial E}\right) G_{T=0}(E). \tag{2.58}$$

For zero temperature, the derivative becomes a δ-function and only states lying on the Fermi circle contribute to the conductance. However, for $T \neq 0$ this function smoothly decreases going away the Fermi level. The width of the peak is approximately $\sim 4k_BT$, k_B being the Boltzmann constant, so that with increasing temperature averaging becomes more effective. The results for nonzero temperature are shown in Figure 2.8. Using Equation 2.43 one can calculate the energy difference between the subbands

$$\epsilon_{n+1} - \epsilon_n = \frac{2n+1}{n^2} E_F, \tag{2.59}$$

and resonances

$$\epsilon_{n,\ell+1} - \epsilon_{n,\ell} = \left(\frac{\lambda_F}{2d}\right)^2 (2\ell+1) E_F. \tag{2.60}$$

Finite temperature averaging washes out the corresponding features whenever $4k_BT$ is equal to or larger than these energy differences. Two observations are in order: First, the distortion of the step structure with increasing temperature is independent of the length of the constriction d, and secondly the resonances disappear more easily for longer constrictions compared to the shorter ones. The first observation, in fact, is not correct in the more general sense, since for short constrictions tunneling is important and even without temperature effects the steps are smooth. This, together with the second observation implies that the QPC showing quantization of conductance has to be rather long. For example, using[58] $\lambda_F = 42$ nm, the energy differences between the first and second resonances are 1.1 °K and 17.6 °K for constrictions of $d = 420$ nm and 105

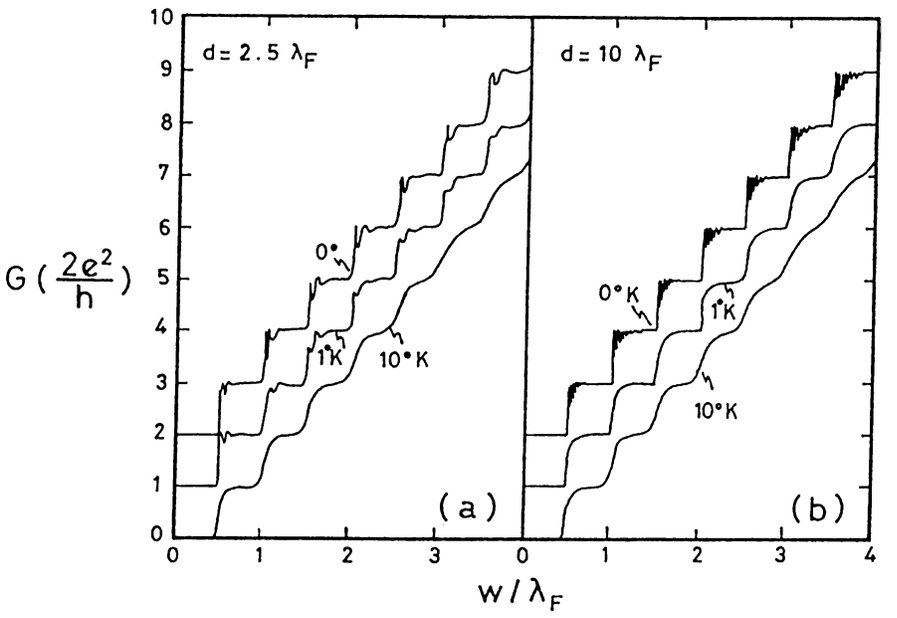

Figure 2.8: Conductance of quantum point contacts at nonzero temperatures The conductance as a function of width for temperatures of 0, 1 and 10 ° K. The length of the constriction is (a) $2.5 \times \lambda_F$ and (b) $10 \times \lambda_F$. The Fermi wavelength is taken to be equal to 42 nm. The curves are offset for clarity. Infinite-well confinement.

nm long, respectively[‡]. Clearly, for temperatures higher than 250 mK these resonances would be washed out for the longer one, compared to 4° K for the shorter one. For resonances with higher ℓ this critical temperature will be higher. On the other hand, as seen in Figure 2.7 the oscillations due to the resonances get smaller with ℓ and thus their effect will be of minor importance. Note, however, that for constrictions longer than $d = 10\ \lambda_F$ the quantization begins to disappear[112] due to the voltage fluctuations[113] and the resulting localization in the constriction. Thus, the quantization of conductance and lack of resonances are possible only for an intermediate range of d. The experiments carried out at sub-Kelvin temperatures[110] displayed some irregular features resembling to theoretically conjectured resonances. Nevertheless, it is clear[84,88,96] that the

[‡]For the same parameters the temperature necessary to wash out the 10th quantized step is found to be 31 °K. For this temperature it is not possible to have ballistic transport. Thus, the distortion of steps with temperature is mainly due to the transition to the diffusive regime.

temperature effects do not suffice to explain the experimental results. In the next section we present another mechanism which supplies a better explanation.

We consider the effects of finite bias for $T = 0$, since we understood the averaging due to the temperature. For a finite bias the electrochemical potential of the left-hand reservoir is $\mu+eV/2$ and that of the right-hand one is $\mu-eV/2$. We assume that the bias does not affect the solutions in the constriction appreciably. Thus the current passing through the constriction is

$$I = \int_{\mu-eV/2}^{\mu+eV/2} dE\ J(E), \tag{2.61}$$

and thus the conductance is

$$G_V = \frac{1}{eV} \int_{\mu-eV/2}^{\mu+eV/2} dE\ G_{V=0}(E). \tag{2.62}$$

Clearly, the effects of finite bias are very similar to those of finite temperature. For long constrictions the resonance structure disappears for small but finite bias, on the contrary for short constrictions the resonances are still present. Increasing the bias further, the steps are washed out and quantization is lost. However, in the experimental studies[58,59] the bias is kept low, $eV < k_BT$, in order to prevent heating by the electrons. Thus the bias effects can not be responsible for the lack of resonances. On the other hand, for finite bias the nonlinear conductance of the QPC becomes an important issue, and was investigated both experimentally[114] and theoretically.[96,115,116]

2.3 Nonuniform Constrictions

2.3.1 Transfer Matrix Method

In order to include the geometrical effects which may play crucial roles in the quantization of conductance one has to solve the Schrödinger equation, Equation 2.3, for a general potential $V(y, z)$. In the literature there are some examples for such solutions for special forms of potential.[80,81,92] We are seeking for a method which is capable of dealing with a general potential. This can not

be achieved fully in analytical form. However, exploiting the analogy with the strictly 1D problems one can devise an approximate method.[88,94] Recalling that a uniform constriction has the 1D analogue as a rectangular barrier, a nonuniform constriction has to be identified with a general barrier potential in the 1D case. In strictly 1D systems there is a widely used approximation scheme for solving problems with general potential profiles, namely the transfer matrix method. The main underlying idea for the transfer matrix method is to divide the space into a number of segments and to assume that the potential is constant in each of these segments. Using the usual boundary matching technique at the interfaces between the segments, the approximate solution is obtained for this piecewise constant potential profile. Increasing the number of segments until the convergence of the solution obtained, the deviation between the exact and approximate solutions can be made negligibly small.

Following the same lines of reasoning, we divide the nonuniform constriction into a large number of segments. In each segment $V_s(z)$ and $V_c(y,z)$ are assumed to be constant. Thus, the solution for the subband wave function in the ith segment are the same as that of a uniform constriction with confining potential $V_c(y,z_i)$ and the zero of energy is shifted by $V_s(z_i)$. Then the subband wave functions in the interval $z_{i-1} < z < z_i$ reads

$$\xi_{nE}(y,z) = \sum_m \phi_m(y,z_i)[e^{i\gamma_m(z_i)z}X_{mn}(z_i) + e^{-i\gamma_m(z_i)z}Y_{mn}(z_i)], \tag{2.63}$$

where the lateral wave function $\phi_n(y,z_i)$ in the ith segment is defined by the equation

$$\left[-\frac{\hbar^2}{2m^*}\frac{d^2}{dy^2} + V_c(y,z_i) + V_s(z_i)\right]\phi_n(y,z_i) = \epsilon_n(z_i)\phi_n(y,z_i), \tag{2.64}$$

with subband energy $\epsilon_n(z_i)$ and propagation constant

$$\gamma_n(z_i) = \sqrt{\frac{2m^*}{\hbar^2}[E - \epsilon_n(z_i)]}, \tag{2.65}$$

being a function of segment index. Hence, the solution for the current carrying

states in the constriction can be written as[‡]

$$\psi_{\mathbf{k}_i}(y,z) = \sum_n [e^{i\gamma_n(z_i)z}\Theta_{n\mathbf{k}_i}(z_i) + e^{-i\gamma_n(z_i)z}\Delta_{n\mathbf{k}_i}(z_i)]\ \phi_n(y,z_i), \qquad z_{i-1} < z < z_i. \tag{2.66}$$

The edges of the constriction are $z_0 = 0$ and $z_N = d$ and the wave function in the 2D EG is expressed in terms of plane waves as it is done before.

Next, we have to find the transfer matrix along the constriction. This is done by matching the wave function and its derivative at the interfaces between the segments. Considering the ith interface at $z = z_i$ $(0 < i < N)$ one finds for the wave function

$$\begin{aligned}\psi_{\mathbf{k}_i}(y,z_i) &= \sum_n [e^{i\gamma_n(z_i)z_i}\Theta_{n\mathbf{k}_i}(z_i) + e^{-i\gamma_n(z_i)z_i}\Delta_{n\mathbf{k}_i}(z_i)]\ \phi_n(y,z_i), \\ &= \sum_n [e^{i\gamma_n(z_{i+1})z_i}\Theta_{n\mathbf{k}_i}(z_{i+1}) + e^{-i\gamma_n(z_{i+1})z_i}\Delta_{n\mathbf{k}_i}(z_{i+1})]\ \phi_n(y,z_{i+1}),\end{aligned} \tag{2.67}$$

and its derivative along z-direction

$$\begin{aligned}\frac{\partial}{\partial z}\psi_{\mathbf{k}_i}(y,z)|_{z=z_i} &= \sum_n i\gamma_n(z_i)[e^{i\gamma_n(z_i)z_i}\Theta_{n\mathbf{k}_i}(z_i) - e^{-i\gamma_n(z_i)z_i}\Delta_{n\mathbf{k}_i}(z_i)]\ \phi_n(y,z_i), \\ &= \sum_n i\gamma_n(z_{i+1})[e^{i\gamma_n(z_{i+1})z_i}\Theta_{n\mathbf{k}_i}(z_{i+1}) - e^{-i\gamma_n(z_{i+1})z_i}\Delta_{n\mathbf{k}_i}(z_{i+1})]\ \phi_n(y,z_{i+1}).\end{aligned} \tag{2.68}$$

Multiplying Equations 2.67 and 2.68 by $\phi_n^*(y,z_i)$ and integrating over y one gets

$$\begin{aligned}&[e^{i\gamma_n(z_i)z_i}\Theta_{n\mathbf{k}_i}(z_i) + e^{-i\gamma_n(z_i)z_i}\Delta_{n\mathbf{k}_i}(z_i)] = \\ &\sum_m S_{nm}(z_i)[e^{i\gamma_m(z_{i+1})z_i}\Theta_{m\mathbf{k}_i}(z_{i+1}) + e^{-i\gamma_m(z_{i+1})z_i}\Delta_{m\mathbf{k}_i}(z_{i+1})],\end{aligned} \tag{2.69}$$

and

$$\begin{aligned}&\gamma_n(z_i)[e^{i\gamma_n(z_i)z_i}\Theta_{n\mathbf{k}_i}(z_i) - e^{-i\gamma_n(z_i)z_i}\Delta_{n\mathbf{k}_i}(z_i)] = \\ &\sum_m S_{nm}(z_i)\gamma_m(z_{i+1})[e^{i\gamma_m(z_{i+1})z_i}\Theta_{m\mathbf{k}_i}(z_{i+1}) - e^{-i\gamma_m(z_{i+1})z_i}\Delta_{m\mathbf{k}_i}(z_{i+1})],\end{aligned} \tag{2.70}$$

[‡]Note that, we combined the coefficients of the uniform constriction wave functions X and Y in the subband wave functions ξ_{nE} and the coefficients of the subband wave functions in the current carrying states $\psi_{\mathbf{k}}$ in order to find the solution for the whole system at a single step.

respectively, where S_{nm} denotes the overlap of the lateral wave functions in the ith and $(i+1)$th segments

$$S_{nm} = \int dy\ \phi_n^*(y, z_i)\ \phi_m(y, z_{i+1}). \tag{2.71}$$

Note that, in each segment the lateral wave functions form an orthonormal complete set. The lateral wave functions in different segments, however, has not such a property if the confining part of the potential $V_c(y,z)$ is different in two segments. Writing Equations 2.69 and 2.70 in a matrix form, the transfer matrix for the ith interface can be found as

$$\tilde{\mathbf{T}}_{i,i+1} = \frac{1}{2}\begin{bmatrix} e^{-i\tilde{\Gamma}(z_i)z_i}\{\tilde{\mathbf{S}}_{i,i+1} + \tilde{\Gamma}(z_i)^{-1}\tilde{\mathbf{S}}_{i,i+1}\tilde{\Gamma}(z_{i+1})\}e^{i\tilde{\Gamma}(z_{i+1})z_i} & \\ e^{i\tilde{\Gamma}(z_i)z_i}\{\tilde{\mathbf{S}}_{i,i+1} - \tilde{\Gamma}(z_i)^{-1}\tilde{\mathbf{S}}_{i,i+1}\tilde{\Gamma}(z_{i+1})\}e^{i\tilde{\Gamma}(z_{i+1})z_i} & \\ & e^{-i\tilde{\Gamma}(z_i)z_i}\{\tilde{\mathbf{S}}_{i,i+1} - \tilde{\Gamma}(z_i)^{-1}\tilde{\mathbf{S}}_{i,i+1}\tilde{\Gamma}(z_{i+1})\}e^{-i\tilde{\Gamma}(z_{i+1})z_i} \\ & e^{i\tilde{\Gamma}(z_i)z_i}\{\tilde{\mathbf{S}}_{i,i+1} + \tilde{\Gamma}(z_i)^{-1}\tilde{\mathbf{S}}_{i,i+1}\tilde{\Gamma}(z_{i+1})\}e^{-i\tilde{\Gamma}(z_{i+1})z_i} \end{bmatrix}, \tag{2.72}$$

which connects the solution in the ith segment to that in the $(i+1)$th segment as

$$\begin{bmatrix} \tilde{\Theta}_{\mathbf{k}_i}(z_i) \\ \tilde{\Delta}_{\mathbf{k}_i}(z_i) \end{bmatrix} = \tilde{\mathbf{T}}_{i,i+1}\begin{bmatrix} \tilde{\Theta}_{\mathbf{k}_i}(z_{i+1}) \\ \tilde{\Delta}_{\mathbf{k}_i}(z_{i+1}) \end{bmatrix}. \tag{2.73}$$

In Equation 2.72 the subscripts of $\tilde{\mathbf{S}}$ and $\tilde{\mathbf{T}}$ refer to the segments for which they are calculated, and not to the subband indices of the matrices. The solution in the first and the last segments, in turn, are connected to each other by

$$\begin{bmatrix} \tilde{\Theta}_{\mathbf{k}_i}(z_1) \\ \tilde{\Delta}_{\mathbf{k}_i}(z_1) \end{bmatrix} = \tilde{\mathbf{T}}_{1,N}\begin{bmatrix} \tilde{\Theta}_{\mathbf{k}_i}(z_N) \\ \tilde{\Delta}_{\mathbf{k}_i}(z_N) \end{bmatrix}, \tag{2.74}$$

where $\tilde{\mathbf{T}}_{1,N}$ is the product of all the transfer matrices along the constriction

$$\tilde{\mathbf{T}}_{1,N} = \prod_{i=1}^{N-1} \tilde{\mathbf{T}}_{i,i+1}, \tag{2.75}$$

where the the order of transfer matrices is the same as their physical appearance. The continuity equations at $z = 0$ give

$$(2\pi)^{1/2}\ 2k_o\ \tilde{\Phi}^{\dagger}(\kappa_o, z_0) = \{[\tilde{\mathbf{K}}(z_1) + \tilde{\Gamma}(z_1)]\tilde{\Theta}_{\mathbf{k}_i}(z_1) + [\tilde{\mathbf{K}}(z_1) - \tilde{\Gamma}(z_1)]\tilde{\Delta}_{\mathbf{k}_i}(z_1)\}, \tag{2.76}$$

and those at $z = d$ give

$$0 = \{[\tilde{\mathbf{K}}(z_N) - \tilde{\mathbf{\Gamma}}(z_N)]\ e^{i\tilde{\mathbf{\Gamma}}(z_N)d}\tilde{\mathbf{\Theta}}_{\mathbf{k}_i}(z_N) + [\tilde{\mathbf{K}}(z_N) + \tilde{\mathbf{\Gamma}}(z_N)]\ e^{-i\tilde{\mathbf{\Gamma}}(z_N)d}\tilde{\mathbf{\Delta}}_{\mathbf{k}_i}(z_N)\}. \tag{2.77}$$

In Equation 2.76, $\tilde{\mathbf{\Phi}}(\kappa_o, z_0)$ is the vector of transverse Fourier transforms of the lateral wave function in the first segment. Equations 2.74, 2.76 and 2.77 can be solved simultaneously to obtain the wave function in the Nth segment. The wave function in any other segment can be calculated therefrom by using the transfer matrices. The conductance of the constriction can be calculated by using Equation 2.17 and the wave function given in Equation 2.66. The result is exactly in the same form as Equation 2.40. The segment in which $\tilde{\mathbf{\Gamma}}$, $\tilde{\mathbf{\Theta}}_{\mathbf{k}_i}$ and $\tilde{\mathbf{\Delta}}_{\mathbf{k}_i}$ are calculated does not matter since the current along the constriction is conserved. In what follows we use the transfer matrix method to calculate the conductance for several constriction geometries.

2.3.2 Adiabatic Evolution of States

To analyze the effects of geometry on quantization, we first consider wedge-like and flared constrictions. In Figure 2.9 the conductance for wedge-like constrictions with varying wedge-angle α are shown. Clearly a wedge-like constriction has lower subband energies at the wider side compared to those at the narrower side. Thus, its strictly 1D analogue resembles to a trapezoidal barrier. The tunneling probability through a trapezoidal barrier is larger compared to a rectangular barrier with the same maximum barrier height. Consequently, for a wedge-like constriction, the smaller is α the larger is the contribution of the evanescent states in the conductance. The $\alpha = 0$ case corresponds to the quantum Sharvin limit and $\alpha = 90^o$ to the uniform constriction. An important observation is that for intermediate wedge-angle values ($65^o < \alpha < 85^o$) the steps begin to form out of the oscillations observed for the quantum Sharvin conductance. However, the resonances are not visible although the length of the constriction is long enough. The reason for this is that the relative phase of the interfering waves is no more proportional to the length of the constriction, so effectively

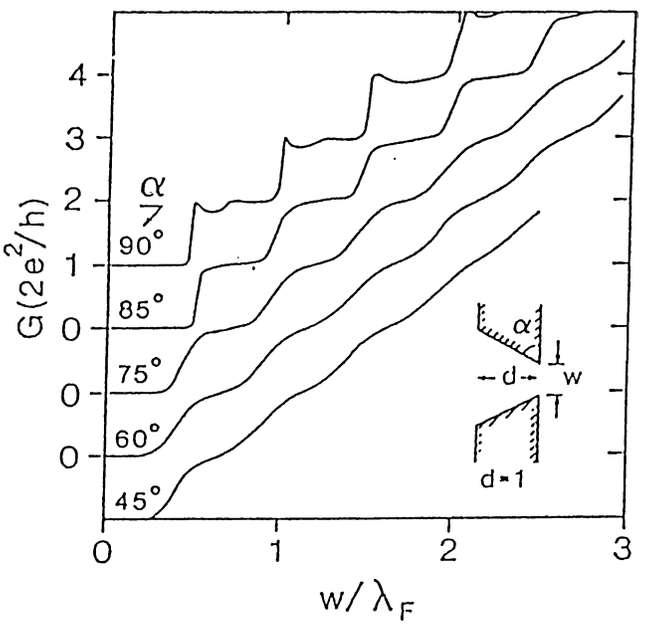

Figure 2.9: Conductance for wedge-like quantum point contact geometries The conductance is given as a function of width. The length of the constriction is λ_F. The curves are offset for clarity. The inset shows the geometry. Infinite-well confinement.

the constriction is shortened. In addition, the reflections at the entrance of the wedge-like constriction are greatly reduced due to the larger width and thus, the amplitude of the oscillations would have been smaller compared to the uniform channel even if they have existed.

The quantization of conductance can be improved by inserting a uniform segment between two wedge-like openings to the 2D EG. For such a structure the uniform part is the narrowest region and the tunneling probability is smaller compared to a single wedge-like constriction. Therefore, the steps become sharper and the plateaus become flatter as shown in Figure 2.10. Note that for small α and long wedge-like openings the resonance structure is still present. This is due to the strong intersubband coupling in the wedges and the resulting nonnegligible reflections. Increasing α, the naive expectation is to have more pronounced resonance effects, since the structure resembles a uniform constriction. However,

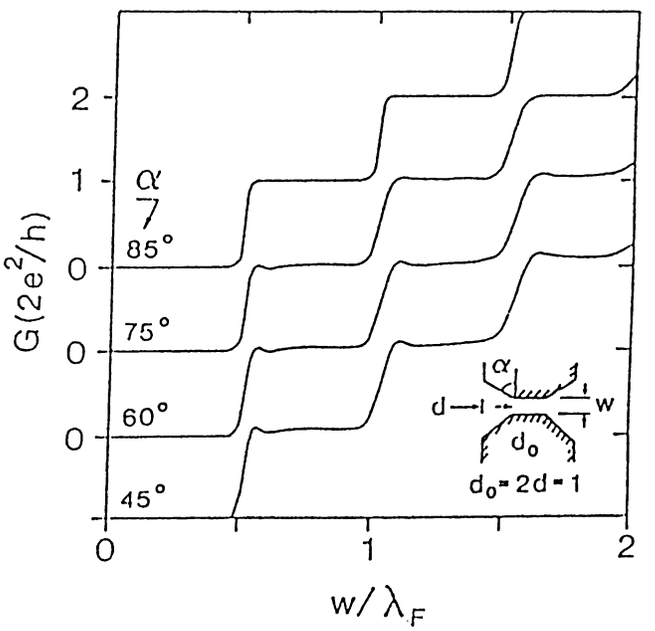

Figure 2.10: Conductance for tapered quantum point contact geometries The conductance is given as a function of width. The length of the constriction is $2 \times \lambda_F$ and $\lambda_F/2$ sections on both edges are tapered. The inset shows the geometry. The curves are offset for clarity. Infinite-well confinement.

the result is just to the contrary, namely the resonance structure disappears. Analyzing G_t defined in Section 2.2 for a flared entrance into the uniform channel one finds that increasing the wedge angle α, the steps in conductance get sharper and plateaus get flatter. The flaring, however, becomes ineffective for $\alpha \rightarrow 90°$, and the steps approach to those obtained for the uniform constriction shown in Figure 2.3. This enhancement in transmission with increasing α is due to the adiabatic evolution[81] of states through the QPC. The adiabaticity of transport implies that the *slow* lateral coordinate can be eliminated from the Schrödinger equation, Equation 2.3, so that the problem reduces to a 1D differential equation for the *fast* longitudinal coordinate. The necessary condition to achieve adiabatic transport is the slow variation of the width (or the confining potential) on length scale of electron wavelength. For such a system the subband wave functions belonging to different subbands are completely decoupled and each satisfies its

own *effective* Schrödinger equation

$$\left[-\frac{\hbar^2}{2m^*}\frac{\partial^2}{\partial z^2} + V_s(z) + V_{n,eff}(z)\right]\zeta_{nE}(z) = E\zeta_{nE}(z), \tag{2.78}$$

where $\xi_{nE}(y,z) = \zeta_{nE}(z)\phi_n(y,z)$, and $\phi_n(y,z)$ is the lateral wave function corresponding to $V_c(y,z)$. The effective potential $V_{n,eff}(z)$ is equal to the subband energy $\epsilon_n(z)$ plus some corrections.[89] In short, the adiabatic approximation verifies the analogy between the QPC and strictly 1D problems, and shows that for certain cases the analogy becomes an identity. It is shown[81] that adiabatic transport leads to quantization of conductance without any resonance effects. The quantization originates from the fact that the transmission through the system is determined by the transmission at the narrowest portion, which is a uniform constriction of width w. The lack of the resonances is a result of the suppression of intersubband scattering and intrasubband reflections due to the smooth variation of the effective potential. However, the justification of the adiabatic approximation as used by Glazman and coworkers[81] have some strict conditions. The local adiabaticity, i.e., the applicability of the above formalism around the narrowest part of the constriction, requires $R \gg w_o$, R and w_o being the radius of curvature and width of the constriction at the narrowest point, respectively. In order to adiabatic picture to be valid throughout the device, i.e., for global adiabaticity, the length of the device L has to be large enough so that $L \gg \sqrt{w_o R}$. The suppression of reflection and intrasubband scattering, on the other hand, is possible only for $k_F R \gg 1$. It is not possible to have all these conditions satisfied for realistic structures, however. Later Yacoby and Imry[89] showed that the adiabaticity effects are important even for finite constrictions which have abrupt junctions to the 2D EG and the corrections due to this abrupt junctions are exponentially small. They found that a nonadiabatic opening has small effects whenever $w_o \ll w$, w being the width of the constriction at the position of the nonadiabatic opening. On the other hand, smoothly varying but nonadiabatic width variation does not change the conductance appreciably for $N_{w_o} < N_w$, N being the number of subbands below E_F. Tekman and Ciraci[94] using the transfer matrix method described above

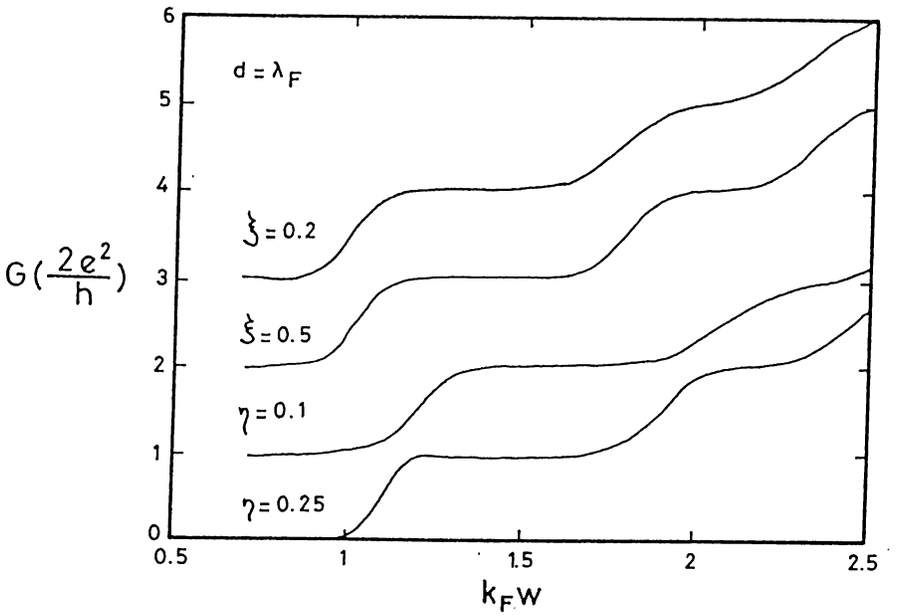

Figure 2.11: Conductance of quantum point contacts with saddle point potentials
The conductance as a function of width for a constriction of length λ_F. The potential and width profiles are as given in Equations 2.79 and 2.80. The curves are offset for clarity. Parabolic confinement.

calculated the conductance for finite length constrictions with smoothly varying width. They found that even the resonances are still present the amplitude of these resonances are $\lesssim 0.01 \times 2e^2/h$, which is very close to the ideal step structure found for infinite electron waveguides. Recently a consensus about the importance of adiabatic evolution of states for the experimentally relevant structures and lack of resonances was established.

2.3.3 Transmission through a Saddle Point

Earlier, Ciraci and Tekman used the saddle point structure calculated from the self-consistent field calculations to investigate the transition from tunneling to ballistic regime in STM. Recently Büttiker[117] analyzed the effect of the saddle point on the quantization. Using the adiabatic approximation he showed that

the conductance is quantized under certain conditions. It is now established that the quantization of conductance can be affected significantly depending on the form of the saddle potential. We use the transfer matrix approach described in this section to investigate the effects of a saddle point potential for a finite length QPC. This is an interesting example since for the realistic system the electron density is smaller just below the split-gate than it is in the 2D EG. This indicates that the existence of a saddle point potential for realistic QPC. According to the adiabatic picture, the width of the constriction acts as an effective potential, so that there are two ways of analyzing the effects of the saddle point. The first one is to include a longitudinal potential

$$V_s(z) = \eta \; E_F \; \left(1 - \left[\frac{z - d/2}{d/2}\right]^2\right), \tag{2.79}$$

which yields a potential barrier of height ηE_F at the center of the constriction. Then, the Schrödinger equation, Equation 2.3, is separable and the subband wave functions are products of the 1D solution for this potential and the lateral wave function. The limit $d \rightarrow \infty$ corresponds to the system analyzed by Büttiker.[117] Our analysis, however, include the effect of the nonadiabatic opening to the 2D EG. On the other hand, changing the width of the constriction in the form

$$w(z) = w_o + \frac{\xi}{k_F} \; \left(\frac{z - d/2}{d/2}\right)^2, \tag{2.80}$$

produces an effective potential barrier which has the saddle point structure. For this case the Schrödinger equation is not separable and use of the transfer matrix method is necessary. The results obtained by using both kinds of constrictions are shown in Figure 2.11. Comparing these conductance curves with that given in Figure 2.6 one sees that the saddle point potential leads to adiabatic evolution of states so that the resonance structure is suppressed. However, as η gets larger the real potential starts to dominate the effective potential and the steps between the plateaus become wider and the quantization disappears. Similarly, as ξ gets larger the length of the effective potential barrier gets smaller so that the evanescent states become more effective, and yields narrowing of the plateaus and distortion of quantization.

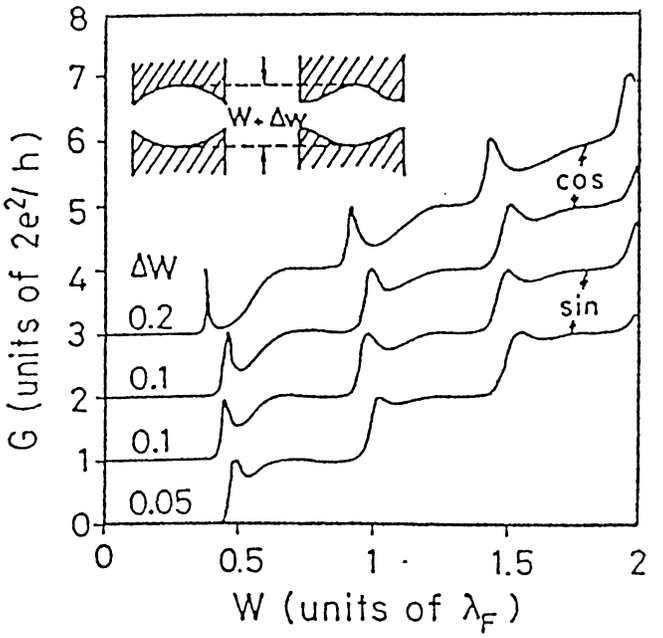

Figure 2.12: Conductance for quantum point contacts with resonant tunneling features

Conductance as a function of width for constrictions of length λ_F. The width of the constrictions are varied by sine- and cosine-modulation. The curves are offset for clarity. The inset shows the geometry used. Infinite-well confinement.

2.3.4 Resonant Tunneling

According to the adiabatic picture of ballistic transport described above, the subband energy $\epsilon_n(z)$ acts as a 1D effective potential. Thus by widening the constriction at the center $z = d/2$, it should be possible to form a cavity (effective potential of which resembles to a quantum well) along the channel. A similar effect can be observed when the longitudinal potential $V_s(z)$ has a minimum at the center of the constriction. Thus, the constriction effectively acts like a quantum well between two potential barriers for certain QPC configurations. Accordingly, such a structure leads to formation of quasi-0D states which are bound to the cavity and give rise to resonant tunneling effects for the finite QPC. Note that, in order to observe resonant tunneling effects it is not necessary to have adiabatic transport in the system. For finite systems with the properties

discussed above it is possible[94,118] observe peaks in the conductance curve, just below opening of a new channel of propagation. Exploiting the analogy between the quasi-1D systems and strictly 1D problems it is concluded that these resonant tunneling geometries are analogous to the double-barrier resonant tunneling (DBRT) structures obtained by *band-gap engineering* methods. Namely, in addition to ordinary tunneling through the channel, there is a resonance event taking place as evidenced from the peaks in the conductance curve. For a certain width of the constriction, resonance takes place and unity transmission through the channel is achieved. Destroying the resonance condition (by increasing the width, thus changing the position of the quasi-0D state with respect to E_F) the conductance rapidly drops to its ordinary tunneling value, however.

Some typical examples of such electron cavities are given in Figure 2.12. The width of the constriction is varied using sine- and cosine-modulation with the amplitude Δw. We found that the sine-modulation yields relatively smaller widths w_{res} (at which resonant tunneling occurs), and broader peaks as compared to those for the cosine-modulation. This can be explained by using the DBRT analogy. Since the sine-modulation is represented by a DBRT structure with a relatively wider well, but narrower barriers as compared to those of the cosine-modulation, the resonance energies are relatively farther from the top of the barrier (or equivalently from the quantized steps in G) and resonance peaks are relatively broader for the sine-modulation. That the DBRT analogy is valid and thus the peaks near the steps are related to the resonant tunneling is verified[119] by calculating the quasi-0D states for an infinite constriction with a widening of the same form as the finite constriction above and by comparing the positions of these with the resonance positions w_r. A detailed analysis of resonant tunneling is presented in Section 2.5.

2.3.5 Effects of the Roughness of the Channel

Finally, we investigate the effect of the roughness along the constriction, which is closely related to the quality of the split-gate. The roughness is simulated[94] by a random modulation of the width. At each step $\delta d = d/N$ along the

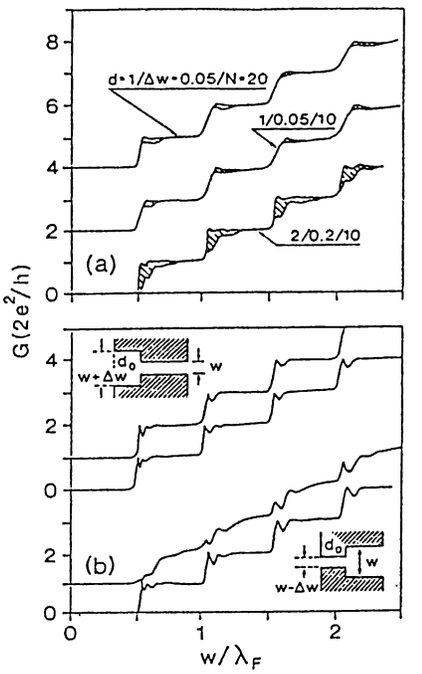

Figure 2.13: Conductance for quantum point contacts with nonideal geometries (a) Conductance as a function of average width for rough channel, (b) conductance as a function of width when an obstacle is present at the entrance. The length of the constrictions is λ_F and $2 \times \lambda_F$ in (a) and (b), respectively. The curves are offset for clarity. The inset shows the geometry. Infinite-well confinement.

constriction w is varied by $x\Delta w$, where the value of x $(0 \leq x \leq 1)$ is random. This way a histogram profile $\delta w(z)$ is superimposed over the uniform width w. The conductance is then calculated for 25 different $\delta w(z)$ profiles, and is traced on the same plot with respect to the average width $w_{av} = w + d^{-1} \int_0^d \delta w(z) dz$. In Figure 2.13(a) the conductance $G(w_{av})$ for a given set of $(d, \Delta w, N)$ varies in the shaded region as a function of the profile. The width of the shaded region increases with increasing Δw. For small Δw, the positions of the resonance peaks are maintained. Nevertheless, the conductance values at the resonances and anti-resonances are affected because of the scattering from the roughness. Note that, it was shown[120] that even when the lithographic form of the split-gate has sharp features, the electrostatic depletion region has smooth edges. Thus, our model with abrupt variation of the width causes large distortions as compared to the

realistic ones. Still, the roughness may lead to important consequences such as diffusive reflections from the walls of the constriction as found experimentally.[121]

Two other forms of the constriction are obtained by putting an obstacle of length d_o at the entrance of a uniform constriction, the width at the obstacle being relatively larger or smaller ($\pm\Delta w$). For the first case, since the states in the obstacle region can match the 2D EG states to those of the uniform constriction, the conductance is not affected in any essential manner. In contrast, the narrower obstacle lacks appropriate states, which match the uniform constriction to the 2D EG. Since the opening of the channels are shifted by $+\Delta w$, the sharp step structure is disturbed and the flat plateaus disappeared with increasing Δw. We notice that even for $d_o = 0.05\lambda_F$ a small reduction of w at the entrance give rise to drastic deviations in conductance as compared to that of the uniform constriction. If such an obstacle is put in the constriction near the center its effect is not as important as that of at the entrance, however.

2.4 Elastic Scattering by an Impurity in the Constriction

2.4.1 Theory

Up to this point, we considered the geometrical effects on the transport through the QPC. However, there are several reasons for other elastic scattering processes to take place, such as the potential fluctuations due to the random distribution of the dopants,[113] unintentional doping[122] near the 2D EG and surface roughness at the heterojunction. In this section we analyze the effects of the elastic scattering on the quantization of conductance. The effects of elastic scattering by an impurity in the channel were investigated by the Tekman and Ciraci[98] using Green's function technique[‡]. It is also possible to obtain the solution for an arbitrary scattering potential by using the transfer matrix method described in Section 2.3 as well.

[‡]See Appendix 2.A.

Tekman and Ciraci[98] showed that it is possible to get analytical solutions for certain scattering potentials. These are important for a simple understanding of the underlying physics and to shed light on the properties of the system under consideration. Here we present the solution to the particular problem described in Reference 98 by exploiting the analogy between the quasi-1D and strictly 1D systems. To this end, the scatterer is simulated by the model potential

$$V_I(y,z) = \frac{\hbar^2\beta}{m^*}\delta(z-z_I)v(y), \tag{2.81}$$

where (y_I, z_I) is the position of the scatterer, $|\beta|$ is its strength and $v(y)$ is a function with a peak at $y = y_I$. The impurity is repulsive for $\beta > 0$ and attractive for $\beta < 0$. For the strictly 1D case the solution is obtained by using the characteristic discontinuity of the derivative of the wave function due to the presence of the δ-function potential. In what follows, we apply the same procedure to the quasi-1D case using a uniform constriction in order to make the formulation tractable.

For $z \neq z_I$ the potential is just the uniform confining potential $V_c(y)$ and the solutions of the Schrödinger equation are the linear combinations of the subband wave functions for the uniform constriction. Note that, the presence of the impurity yields a finite probability for backscattering. Hence, the subband wave function can be written as

$$\begin{aligned} \xi_{nE}(y,z) &= e^{i\gamma_n z}\phi_n(y) + \sum_m R_{mn}e^{-i\gamma_m z}\phi_m(y), \qquad z < z_I, \\ &= \sum_m T_{mn}e^{i\gamma_m z}\phi_m(y), \qquad z_I < z. \end{aligned} \tag{2.82}$$

Here $|R_{mn}|^2$ and $|T_{mn}|^2$ are the reflection and transmission probabilities in the mth subband for a incident wave in the nth subband from left. Writing the Schrödinger equation explicitly that ξ_{nE} satisfies

$$\left[-\frac{\partial^2}{\partial z^2} - \frac{\partial^2}{\partial y^2} + \frac{2m^*}{\hbar^2}V_c(y) + 2\beta\delta(z-z_I)v(y)\right]\xi_{nE}(y,z) = k_E^2\xi_{nE}(y,z), \tag{2.83}$$

one observes that the subband wave function and its first derivative along z-direction have to be discontinuous. Continuity of the wave function implies

$$e^{i\gamma_n z_I}\phi_n(y) + \sum_m R_{mn}e^{-i\gamma_m z_I}\phi_m(y) = \sum_m T_{mn}e^{i\gamma_m z_I}\phi_m(y). \tag{2.84}$$

Integrating Equation 2.83 between z_I^- and z_I^+ one gets

$$-\frac{\partial}{\partial z}\xi_{nE}(y,z)|_{z_I^-}^{z_I^+} + 2\beta v(y)\xi_{nE}(y,z_I) = 0, \tag{2.85}$$

using the continuity of the wave function itself. Inserting Equation 2.82 in Equation 2.85 one finds

$$\begin{aligned} i\gamma_n e^{i\gamma_n z_I}\phi_n(y) - \textstyle\sum_m R_{mn} i\gamma_m e^{-i\gamma_m z_I}\phi_m(y) - \sum_m T_{mn} i\gamma_m e^{i\gamma_m z_I}\phi_m(y) \\ +2\beta v(y)\textstyle\sum_m T_{mn} e^{i\gamma_m z_I}\phi_m(y) = 0. \end{aligned} \tag{2.86}$$

Multiplying Equations 2.84 and 2.85 by $\phi_n^*(y)$ and integrating it over y the corresponding matrix equations are obtained as

$$e^{i\tilde{\Gamma}z_I} + e^{-i\tilde{\Gamma}z_I}\tilde{\mathbf{R}} = e^{i\tilde{\Gamma}z_I}\tilde{\mathbf{T}}, \tag{2.87}$$

and

$$i\tilde{\Gamma}e^{i\tilde{\Gamma}z_I} - i\tilde{\Gamma}e^{-i\tilde{\Gamma}z_I}\tilde{\mathbf{R}} - i\tilde{\Gamma}e^{i\tilde{\Gamma}z_I}\tilde{\mathbf{T}} + 2\tilde{\mathbf{u}}e^{i\tilde{\Gamma}z_I}\tilde{\mathbf{T}} = 0, \tag{2.88}$$

where $\tilde{\mathbf{u}}$ is the scattering potential matrix with elements

$$u_{nm} = \beta \int dy\ \phi_n^*(y)v(y)\phi_m(y). \tag{2.89}$$

The solution for ξ_{nE} can easily be found by multiplying Equation 2.87 by $i\tilde{\Gamma}$ from left and adding it to Equation 2.88 one gets

$$\tilde{\mathbf{T}} = e^{-i\tilde{\Gamma}z_I}(\tilde{\Gamma} + i\tilde{\mathbf{u}})^{-1}\tilde{\Gamma}e^{i\tilde{\Gamma}z_I}, \tag{2.90}$$

and from Equation 2.87

$$\tilde{\mathbf{R}} = -ie^{i\tilde{\Gamma}z_I}(\tilde{\Gamma} + i\tilde{\mathbf{u}})^{-1}\tilde{\mathbf{u}}e^{i\tilde{\Gamma}z_I}. \tag{2.91}$$

Equations 2.90 and 2.91 have the analogues in the strictly 1D case as the transmission $[t = k/(k + i\beta)]$ and reflection $[r = -i\beta/(k + i\beta)]$ probability amplitudes, respectively, for a δ-function potential located at the origin. It can be shown that the subband wave function $\bar{\xi}_{nE}$ has a similar form with the reflection and transmission probability amplitudes given by $\bar{T}_{mn} = (T_{mn})^*$ and $\bar{R}_{mn} = (R_{mn})^*$.

The conductance of the infinite channel can be found by using the well known Landauer formula for two-terminal measurements

$$G_\infty = \frac{2e^2}{h}\mathrm{Tr}\{\tilde{\mathbf{T}}_p^\dagger \tilde{\mathbf{T}}_p\}, \tag{2.92}$$

where $\tilde{\mathbf{T}}_p$ is the submatrix of $\tilde{\mathbf{T}}$ including the occupied subbands only. Although this expression seems to be quite different than that obtained by using the Green's function technique,[98] it can be shown that they yield identical results. The current carrying solutions for the finite constriction system can be found by using the equations corresponding to Equation 2.31 and 2.32

$$(2\pi)^{1/2}\, 2k_o\, \tilde{\mathbf{\Phi}}^\dagger(\kappa_o) = \{[\tilde{\mathbf{K}} + \tilde{\mathbf{\Gamma}}](\tilde{\mathbf{1}} + \tilde{\mathbf{R}})\tilde{\mathbf{\Theta}}_{\mathbf{k}_i} + [\tilde{\mathbf{K}} - \tilde{\mathbf{\Gamma}}]\tilde{\mathbf{T}}^*\tilde{\mathbf{\Delta}}_{\mathbf{k}_i}\}, \tag{2.93}$$

and

$$0 = \{[\tilde{\mathbf{K}} - \tilde{\mathbf{\Gamma}}]\, e^{i\tilde{\mathbf{\Gamma}}d}\tilde{\mathbf{T}}\tilde{\mathbf{\Theta}}_{\mathbf{k}_i} + [\tilde{\mathbf{K}} + \tilde{\mathbf{\Gamma}}]\, e^{-i\tilde{\mathbf{\Gamma}}d}(\tilde{\mathbf{1}} + \tilde{\mathbf{R}}^*)\tilde{\mathbf{\Delta}}_{\mathbf{k}_i}\}, \tag{2.94}$$

where the wave function is expressed as it is done in Equation 2.4. Solving these two equations simultaneously and inserting the result in Equation 2.18 one can obtain the conductance G_d of the finite constriction with an impurity in it. The final expression is complicated and is not given here.

2.4.2 Results

The variation of G_∞ for an infinite constriction having a single impurity is shown in Figures 2.14 and 2.15. As seen, the ideal quantization is distorted in the presence of the scatterer and some novel effects arise due to scattering. If the potential of the impurity is weak [e.g., $|\beta| \lesssim 0.5\ k_F$ for $q = 10\ \lambda_F^{-1}$ in Figure 2.14(a)], G_∞ still has a staircase structure with smoothed steps and with plateaus very close to the quantized values given in Equation 1.30. Another observation is that for weak scatterers the sign of the potential does not have a pronounced effect on the conductance. This result is in compliance with the first-order Born approximation, since the lowest order correction to the conductance is proportional to β^2 in the perturbative treatment of the impurity. Therefore both repulsive and attractive impurities have the same effects on the transport.

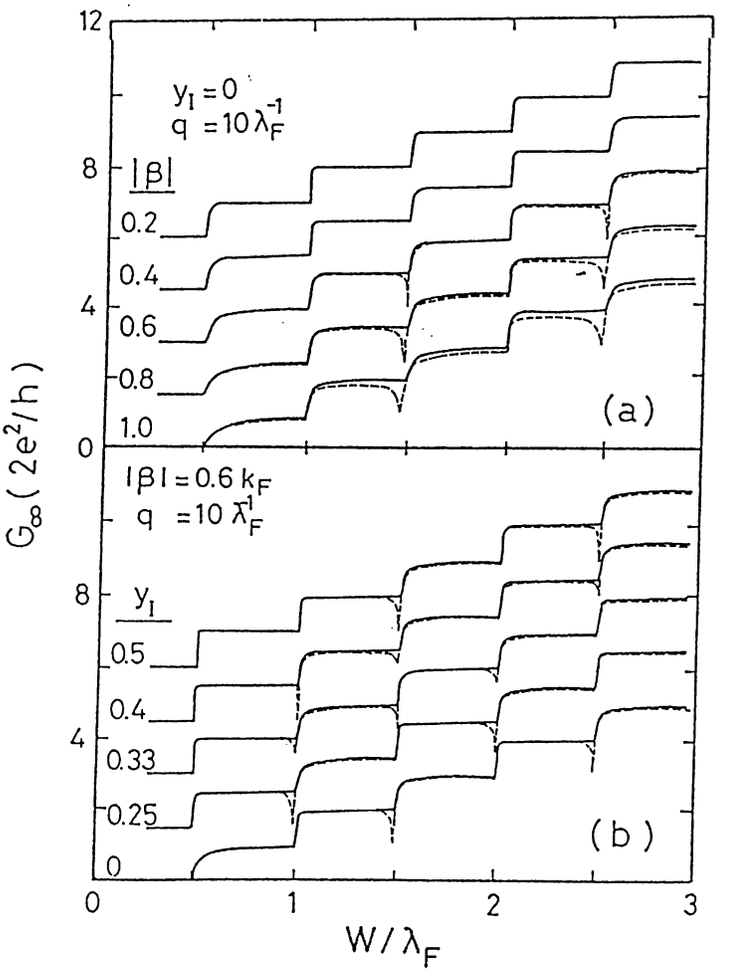

Figure 2.14: Conductance of infinite channel with a single laterally confined impurity
The conductance as a function of width. (a) $y_I = 0$ and $q = 10\ \lambda_F^{-1}$, the strength $|\beta|$ varying in units of k_F., (b) $|\beta| = 0.6\ k_F$ and $q = 10\ \lambda_F^{-1}$, the position y_I varying in units of λ_F. Solid and dashed lines are for repulsive and attractive impurities, respectively. The curves are offset for clarity. Infinite-well confinement.

In order the Born approximation to be valid and thus only a single scattering event to take place the velocity or equivalently the wave vector of electrons has to be large. In the quasi-1D system under investigation the related wave vector is the propagation constant γ_j and is equal to zero whenever a new subband dips

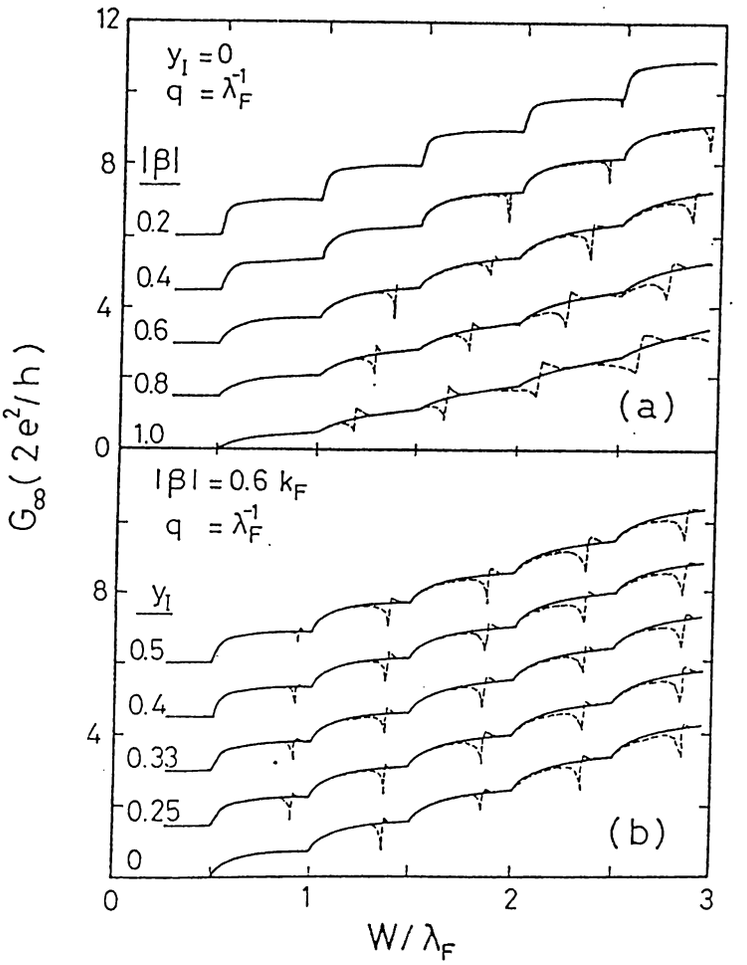

Figure 2.15: Conductance of infinite channel with a single laterally spread impurity
The conductance as a function of width. (a) $y_I = 0$ and $q = \lambda_F^{-1}$, the strength $|\beta|$ varying in units of k_F., (b) $|\beta| = 0.6\ k_F$ and $q = \lambda_F^{-1}$, the position y_I varying in units of λ_F. Solid and dashed lines are for repulsive and attractive impurities, respectively. The curves are offset for clarity. Infinite-well confinement.

the Fermi level, i.e. $w = j\lambda_F/2$. Thus, the Born approximation fails for w values just above $j\lambda_F/2$ and it is necessary to include the multiple scattering events.

For relatively stronger impurities [$|\beta| \lesssim k_F$ for $q = 10\ \lambda_F^{-1}$ in Figure 2.14(a)] not only the steps are smoothed, but also the plateaus exhibit deviations from

the quantized values. The most remarkable effect observed in this range of β is the difference between the attractive and repulsive scatterers. Also for strong scatterers the first-order Born approximation begins to fail for the whole range of w. As seen, for attractive impurities the dips in the G versus w curves appear below the steps. The conductances at these minima are approximately equal to the quantized conductance of the lower plateau for laterally confined impurities, and there are sharp rises to the next quantized value above these dips. One important point we notice is that, dips do not occur below all of the steps. To analyze this, we calculate G_∞ for different transverse positions y_I of the impurity as shown in Figure 2.14(b), and consider u_{jj} given by Equation 2.89. For a laterally confined impurity u_{jj} is approximately proportional to $|\phi_j(y_I)|^2$. To a first approximation, the effect of the impurity is large on the jth plateau when $|\phi_j(y_I)|^2$ is a maximum, but is small when $|\phi_j(y_I)|^2$ is zero. For example, for $y_I = 0$ the deviations from the quantized values will be large on the odd numbered plateaus and small on the even numbered ones. On the other hand, the size, width and existence of the dips below the jth step for the attractive impurities are determined by magnitudes of u_{ij} for $i < j$. Analyzing these dips in detail we find that they originate from the enhancement of backscattering due to the intersubband interaction. For the strictly 1D problem total backscattering is not allowed since the boundary condition at $z = z_I$ for the derivative of the wave function can not be satisfied. For the quasi-1D case the subbands may be coupled in the presence of the impurity. Therefore, the total backscattering can occur in a subband by inclusion of the evanescent states. Since the first-order Born approximation employs the equivalent 1D problem for each subband, the dips can not be obtained perturbatively. The backscattering effect is visible in Figure 2.15(a). For an impurity positioned at the center of the channel even and odd numbered subbands are completely decoupled. Therefore, there is not a dip below the second step, and the dip below the third step is due to the enhanced backscattering in the first subband caused by the evanescent third subband state. In the presence of a large number of impurities all the subbands are mixed and it is possible to observe dips below all of the steps.

For laterally spread impurity potentials with small q the deviations from the quantized steps [Figure 2.15(a)] are enhanced compared to those with large q. For example, the dips do not have quantized conductance values. This is due to the large integrated strength $\sim \beta/q$. Note that $\tilde{u}$ given in Equation 2.89 is determined by this integrated strength and not solely by the strength β. Another observation is that for attractive impurity potentials the dips are shifted to values of w which are smaller than multiples of $\lambda_F/2$ and appear together with peaks. Since the impurity potential abuts a wide region of the constriction, the wave function evaluated at y_I can not give an idea about the effect of the scatterers. Thus, although the deviations from the quantized plateaus vary with y_I [Figure 2.15(b)], this effect is not as drastic as it was for large q.

Comparing these results with those obtained earlier[87,99–101,123] it is concluded that the present model potential is more appropriate to analyze the transport in a ballistic channel with a single impurity. Although the dips were also found in these studies, all of the steps were alike in the results given by Chu and Sorbello[99] since the position of the scatterer is chosen to yield coupling of all of the subbands. In addition to that, their approach does not allow to vary the strength and the integrated strength independently. Therefore the results presented in Figure 2.15 are unique to the present study. Another important advantage of the present approach is that it enables the control of the intersubband coupling. For large values of q the scatterer looks like a δ-function, which enhances the intersubband interactions. In contrast, for small q the potential becomes flat in the lateral direction and the intersubband interaction vanishes, yielding the dips to disappear.

Having discussed elastic scattering due to a single impurity in an infinite constriction, we next consider the situation in a QPC of finite length d. The results are summarized in Figure 2.16. As for the impurity-free constriction the main effect of the finite length is to smooth out the sharp changes in G_∞ (or its first derivative) by the inclusion of evanescent states. This effect is of major importance for short constrictions ($d \lesssim \lambda_F$). As mentioned in Section 2.2, for longer constrictions the effect of evanescent states decreases, but a new feature

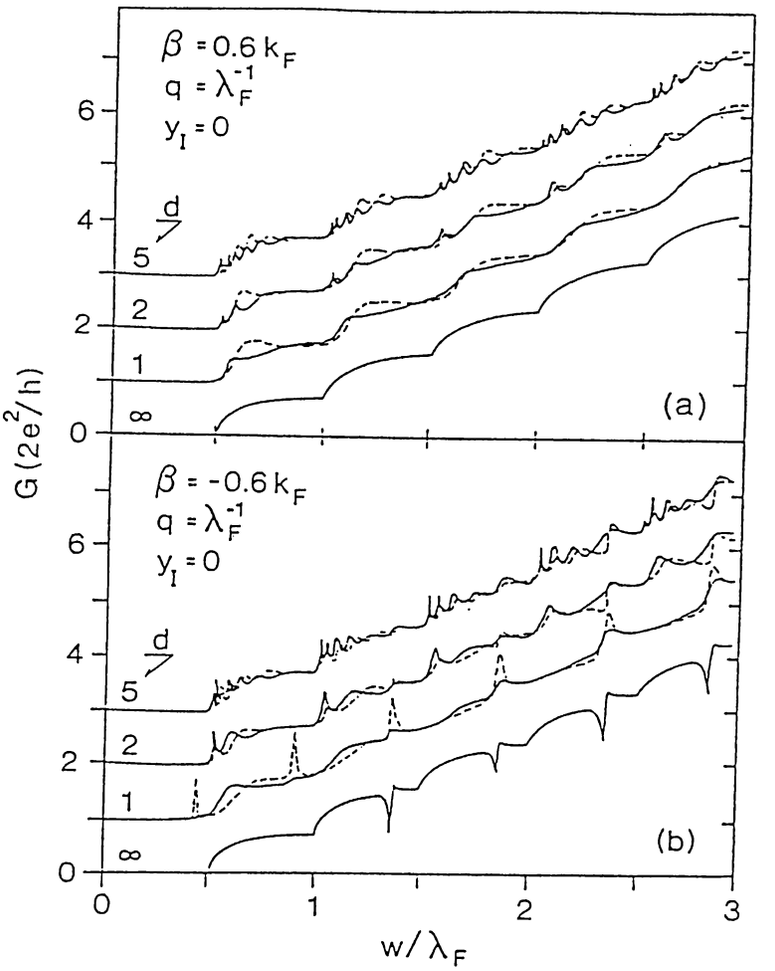

Figure 2.16: Conductance of quantum point contacts with a single impurity The conductance as a function of width. $y_I = 0$, $q = \lambda_F^{-1}$ and the strength $|\beta|$ is (a) 0.6 k_F., (b) -0.6 k_F. The length of the constriction d is varying in units of λ_F. Solid and dashed lines are for $z_I = 0.2$ λ_F and 0.5 λ_F, respectively. The curves are offset for clarity. Infinite-well confinement.

due to the interference of the left- and right-going waves arise, namely the resonance structure. Since the effects of the elastic scattering are taken into account by using a single impurity, neither a phase breaking due to an inelastic event nor a phase averaging due to a large number of scatterers can take place. In other words, the system we are investigating here is in the quasi-ballistic regime,

which still contains well defined interference effects leading to the resonance structure in Figure 2.16. The dramatic effect of the impurity is revealed by comparing conductances of finite (neglecting the contribution due to tunneling) and infinite constrictions. For an impurity-free channel the conductance of the finite constriction is smaller than that of an infinite constriction (i.e., smaller than the ideal quantized steps) for all w. In contrast, for a constriction with a single impurity G_d may be larger than G_∞. This is a result of the combined scattering from the impurity and the ends of the constriction ($z = 0$ and $z = d$). That is, scattering from the ends may depress the effect of scattering by the impurity.

Clearly, the main features of G_d shown in Figure 2.16, in particular the heights and positions of the resonances and anti-resonances are strongly dependent on the position of the impurity along the z-direction. That is, moving the impurity along the channel will give rise to oscillations in the conductance. The magnitude and period of the oscillations are related to the length and width of the constriction, and the properties of the impurity as well. A similar effect was observed[124] by moving defects in a metallic nanoconstriction. In Figure 2.15 the conductance of the QPC for large q has plateaus which are deviated from the quantized values and this deviation decreases with increasing subband index. This result closely resembles to the experimental observation of Wharam and coworkers.[125] These examples show that it is possible to observe the effects of elastic scattering in the channel. Recently, the effect of localized states on backscattering was verified experimentally[122] by observing the dips in the conductance. Although the constrictions used in this experiment were containing number of impurities, the qualitative features are the same with the subject matter of this section and the arguments used to understand the enhanced backscattering are applicable.

Finally, we wish to point out a novel feature of attractive impurities. For short constrictions with an attractive impurity placed near their center ($z_I = d/2$) the conductance curve G_d has sharp peaks just below the steps [Figure 2.16(b)]. The widths of these peaks gets smaller with increasing d, and for very long constrictions the peaks can not even be resolved. Moving the impurity away from the center of the constriction (by changing either the position of the impurity z_I

or the length of the constriction d) has the same effect. Similarly increasing the strength or integrated strength of the impurity causes the peaks to shift the lower w values. A detailed analysis of these results show that these peaks are associated with resonant tunneling through quasi-0D states bound to the impurity. The properties of this resonant tunneling effect are analogous to those obtainable from the double barrier resonance tunneling structures. Hence, similar to formation of quasi-0D states due to the geometrical effects (local widening of the constriction) in an impurity-free ballistic channel, it is possible to obtain bound states in a constriction in the presence of an attractive impurity potential. A final remark about these resonances is that the peaks appear exactly at the same positions with the peaks above the dips in G_∞. This is due to the presence of two orthogonal solutions, one being a quasi-0D state and the other is the current carrying state with unity transmission. Although G_∞ is calculated by including only the current carrying states, G_d has contributions from both of the above states. Therefore the effects of both quasi-0D state and current carrying states are visible in Figure 2.16. The resonant tunneling effect is usually depressed for laterally confined impurities since the resonance peaks and steps are very close to each other, yielding the overlap of corresponding features in G_d. An important remark is about the difference of the evanescent states leading the dips and peaks. Although the dips are formed as a result of the enhanced backscattering stimulated by the intersubband scattering, the peaks are related to bound (or resonance) states localized around the impurity. By decreasing q it is possible to turn off the intersubband scattering and thus to discard the dips. On the other hand, the resonance states become real bound states in the absence of subband mixing. Thus, while the dips are specific to quasi-1D systems, the peaks are due to resonant tunneling and are achievable for all dimensions. A detailed analysis of resonant tunneling is presented in the following section.

2.5 Quasi-zero-dimensional States and Resonant Tunneling

In Sections 2.3 and 2.4, mention was made on an interesting aspect, namely the resonant tunneling in QPC. In this section we analyze the resonant tunneling effects quantitatively, using the case of an attractive impurity in a uniform constriction. We use the formalism described in Appendix 2.A. Note that, the $\tilde{\mathbf{T}}$-matrix becomes singular whenever the determinant of $(\tilde{\mathbf{\Gamma}}+i\tilde{\mathbf{u}})$ is zero. Clearly, $\tilde{\mathbf{u}}$ is real and $\tilde{\mathbf{\Gamma}}$ is purely imaginary with a positive imaginary part when all of the subband energies are above E_F, i.e., $w < \lambda_F/2$. Thus, for a given negative β this determinant vanishes for $w = w_b(\beta, q, y_I)$. This corresponds to a bound 0D state, for which the wave function decays as $\exp(-|\gamma_b|\,|z-z_I|)$ away from the attractive impurity in an infinite uniform channel of width w_b. For a finite length QPC, this bound state interacts with continuum of states in the 2D EG and RT occurs whenever its energy aligned with E_F. In order to have unity transmission, the interaction of the bound state with the left- and right-hand side 2D EG has to be equal. For a symmetric QPC this is possible only when the impurity is placed at the center of the QPC, i.e., $z_I = d/2$. For $w > \lambda_F/2$ there are propagating subbands so that $\tilde{\mathbf{\Gamma}}$ has both real and imaginary elements. Therefore, true bound states do not exist. In what follows we focus our attention on the real bound state existing below the first subband.

In Figure 2.17(a) the conductance below the first step is shown as the position of the impurity (z_I) is varied. Apparently, below the resonance, the conductance increases exponentially and is independent of z_I. This is the ordinary tunneling through the QPC. The position of the impurity does not affect the conductance in any essential manner. Near the resonance, conductance rises rapidly and approaches its maximum value $G_{res}(\beta, q, y_I, z_I)$ for $w = w_{res}(\beta, q, y_I, z_I) \simeq w_b(\beta, q, y_I)$. For $w > w_{res}$ the conductance decreases and at $w = \lambda_F/2$ the first subband becomes propagating leading to another rise in conductance as shown in Figures 2.12 and 2.16. The important observation is that G_{res} is exactly equal to the quantum of conductance $2e^2/h$ when the impurity is at the center ($z_I = d/2$),

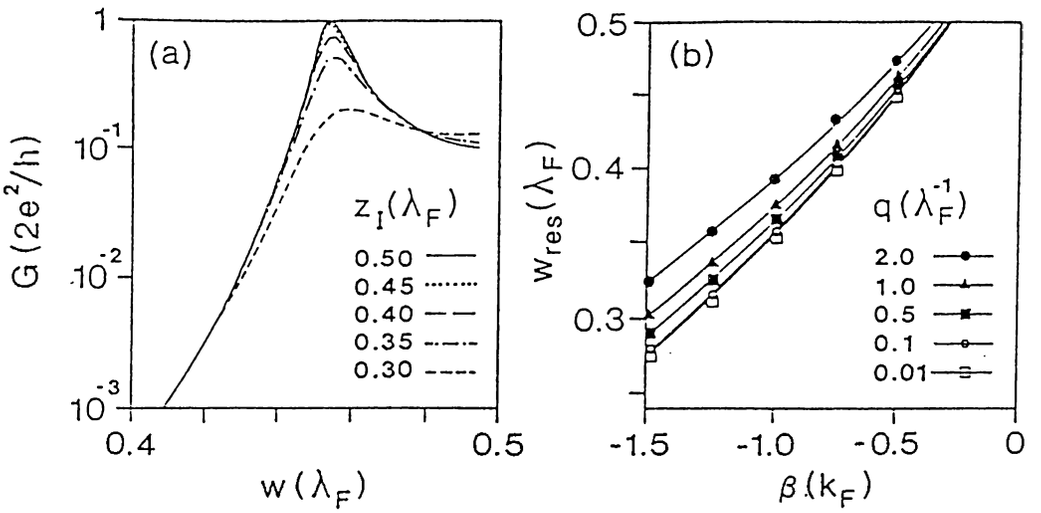

Figure 2.17: Conductance and resonance width for an attractive impurity in the quantum point contact
(a) The conductance as a function of width, the position of impurity being a parameter. The length of the constriction is λ_F and the impurity has $y_I = 0.13\ \lambda_F$, $\beta = -0.5\ k_F$ and $q = 0.5\lambda_F^{-1}$. (b) The resonance width as a function of the strength of the impurity, the lateral extent of the impurity being a parameter. Infinite-well confinement.

but G_{res} decays if the impurity is moved away from the center. This is due to the unbalanced interaction between the 0D state and the left- and right-hand side 2D EG. Also note that, above w_{res} the conductance depends on z_I. This occurs because w_{res} is close to the threshold of propagation for the first subband. Thus, the tunneling probability through the QPC becomes large enough that the boundary conditions at $z = z_I$ determine the conductance to a large extent. Two important quantities extracted from this analysis are w_{res} and the sharpness of resonance. The sharpness is quantified by the full-width at half-maximum (FWHM). For a decreasing resonance peak value, the FWHM increases and w_{res} shifts to larger values. Both of these effects are results of increasing interaction of the bound state with the 2D EG on one side. The larger is this interaction,

the larger is the escape probability, hence the wider is the resonance (the larger is the FWHM). In addition, renormalization of the bound state energy due to the interaction with 2D EG states increases and w_{res} deviates from w_b and a shift to larger w values occurs.

Figure 2.17(b) shows w_{res} as a function of the strength β of the impurity. Note that, the impurity is located at the center ($z_I = d/2$) so that $G_{res} = 2e^2/h$. Naturally the bound states become deeper as the magnitude of the strength is increased. On the other hand, w_{res} increases with increasing q. As pointed out in Section 2.4, the integrated strength of the impurity $\sim \beta/q$ becomes smaller for large q and the bound state gets closer to the propagation threshold, i.e., $w = \lambda_F/2$. For very small q, the impurity potential is almost flat in the transverse dimension so that $\tilde{\mathbf{u}}$ is diagonal. For this case, the bound states are the exact solutions of the 1D problem with a single δ-function scatterer with $w_b = \lambda_F/[4((\beta/k_F)^2 + 1)]^{1/2}$ and binding energy $E_b = -\hbar^2\beta^2/2m^*$ (the energy difference between the bound state and the first subband from which it is split off). For this special case all subbands have real bound states below them. For finite q, that is for nonvanishing intersubband interaction, it is not possible to have a completely localized solution for subbands other than the first one. This leads to the formation of resonance states below the higher lying steps. The actual bound states (lying below the first subband) are used in the analysis of the RT effects in order to focus only on the interaction of the bound state with the 2D EG states, and to eliminate the interactions between the resonance state and the propagating states in the channel. One last observation about w_{res} is that for certain attractive impurities no resonances are observed. In 2D all attractive potentials have bound states, so a w_b can always be identified for attractive impurities. Nevertheless, since the binding energy is extremely small for very small size scatterers, the renormalization overcomes the binding and the bound state does not seem to be split off the subband from which it originates. They can be resolved for longer QPC's, however, due to the reduced renormalization effects.

The effects of the bound state-continuum interactions are analyzed in

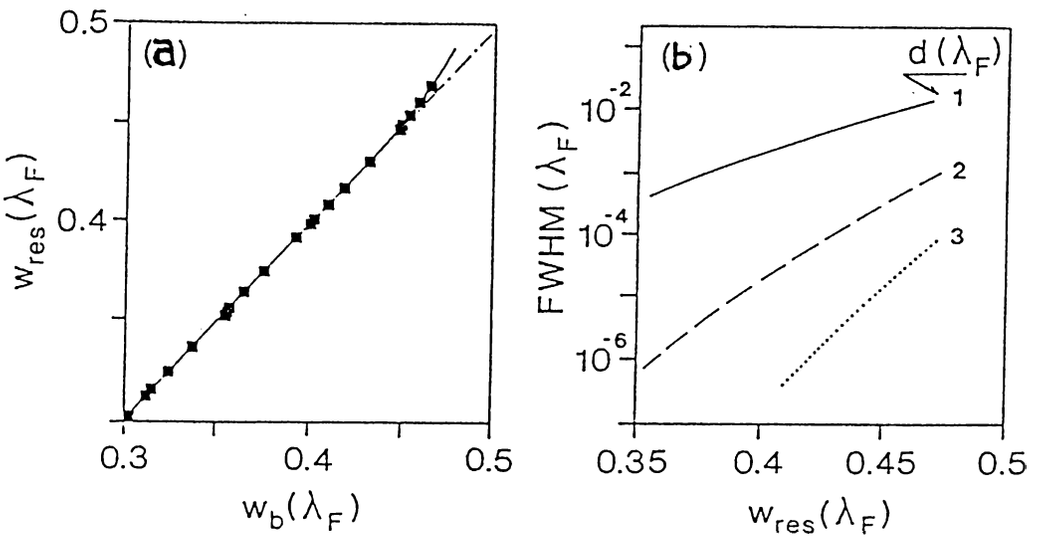

Figure 2.18: Renormalization shift and lifetime for resonance states
(a) The resonance width as a function of bound-state width. The length of the constriction is λ_F, and the impurity is located at the center. (b) Full-width-at-half-maximum as a function of resonance width, the length of the constriction being a parameter. Infinite-well confinement.

Figure 2.18. In Figure 2.18(a) the width w_{res} is compared to w_b which is determined by the zero of the determinant of $(\tilde{\Gamma} + i\tilde{u})$. As long as the bound state is deep, $|\gamma_b|$ is large so the prefactor for hybridization of the bound state with the continuum, $\exp(-|\gamma_b|\ d)$, is negligible. Thus, the resonant transmission occurs exactly at the same width as the bound state, i.e., $w_{res} = w_b$. As the binding energy decreases so does $|\gamma_b|$ and then the hybridization energy may be comparable with E_b. For such a case, the presence of the 2D EG's leads to a shift of the resonance with respect to w_b towards larger w values, that is towards smaller E_b. The deviation of w_{res} from w_b is proportional to $\exp(-|\gamma_b|\ d)$ as expected. The sharpness of resonances are displayed in Figure 2.18(b) for different lengths d. The FWHM decreases as either the binding energy or the length is increased. This is in agreement with the above arguments for the escape rate

and tunneling probability. Since the bound states shown in the figure are far below the propagation threshold, the renormalization effects are not important and the wave function goes as $\exp(-|\gamma_b|\ |z - d/2|)$. Therefore, the FWHM is proportional to $\exp(-c|\gamma_b|\ d)$, which is in agreement with numerical results in Figure 2.18(b) over more than four decades. The exponent c is approximately equal to 1.5, but an analytical calculation of c is quite complicated due to the effects of the quasi-1D nature.

Clearly, the features of resonant tunneling are either in good agreement or analogous to those of the strictly 1D counterpart. Experimental verification of RT events in QPC can be obtained by decreasing the density of electrons (larger λ_F) and by using higher mobility systems. Although observation of Q0D states was reported,[126] it is not clear so far whether these are really RT events since the resistance, but not the conductance curve has peaks.

2.A Impurity Problem and the Lipmann-Schwinger Equation

The eigenstates for a uniform constriction in the presence of a scattering potential $V_I(y,z)$ can be written as

$$\psi_j(y,z) = e^{i\gamma_j z}\phi_j(y) + \int dy'dz'\mathcal{G}(z-z';y,y')V_I(y',z')\psi_j(y',z') \qquad (2.95)$$

The first term on the right hand side represents the incident wave, which is the unperturbed solution for the jth subband and the second term is the scattered part. This expression is the well-known Lippman-Schwinger equation adapted to quasi-1D systems with the retarded Green's function $\mathcal{G}$. The exact solution of Equation 2.95 can be written using the $\mathcal{T}$-operator as

$$\psi_j(y,z) = e^{i\gamma_j z}\phi_j(y) + \sum_n \phi_n(y)\int dk e^{ikz}\tilde{\mathbf{G}}_n(k)\tilde{\mathbf{T}}_{nj}(k), \qquad (2.96)$$

where $\tilde{\mathbf{T}}_{nj}(k)$ is given by

$$\tilde{\mathbf{T}}_{nj}(k) = \tilde{\mathbf{V}}_{nj}(k-\gamma_j) + \sum_m \int dk'\tilde{\mathbf{V}}_{nm}(k-k')\tilde{\mathbf{G}}_m(k')\tilde{\mathbf{T}}_{mj}(k'). \qquad (2.97)$$

Note that the Fourier transforms of $\mathcal{G}$, V_I and $\mathcal{T}$ are matrices $\tilde{\mathbf{G}}$ (diagonal), $\tilde{\mathbf{V}}$ and $\tilde{\mathbf{T}}$, respectively. An element of such matrices are calculated from the integral described by the following expression

$$\tilde{\mathbf{F}}_{nm}(k) = \int dy \phi_n(y) \phi_m(y) \int dz e^{-ikz} f(y,z). \tag{2.98}$$

By solving Equation 2.97 for $\tilde{\mathbf{T}}$ one obtains the solution for the scattering problem for a right-going incident wave in the jth subband. The solution $\bar{\psi}_j(y,z)$ is found similarly for a left-going incident wave. The conductance G_∞ of an infinite constriction is then calculated from the expectation value of the momentum operator,

$$G_\infty = \frac{2e^2}{h} \sum_j^{\epsilon_j < E_F} \frac{1}{\hbar \gamma_j} < \psi_j | \hat{p}_z | \psi_j > . \tag{2.99}$$

The solution of Equation 2.97) for a general potential $V_I(y,z)$ is complicated and may require extensive computations. In order to obtain an analytical solution which leads to a clear picture of the effects of elastic scattering we use the following model potential for a scatterer located at (y_I, z_I)

$$V_I(y,z) = \frac{\hbar^2 \beta}{m^*} \exp(-q|y - y_I|)\ \delta(z - z_I), \tag{2.100}$$

which is δ-function in the z-direction, and has the exponentially decaying form in the y-direction with a decay length of q^{-1}. The strength of this potential is set by the magnitude of β, which may be both attractive ($\beta < 0$) and repulsive ($\beta > 0$). For this form of the potential, Equation 2.97 is exactly solvable and the $\tilde{\mathbf{T}}$-matrix is given by

$$\tilde{\mathbf{T}}_{mj}(k) = (2\pi)^{-1/2}\ e^{-i(k-\gamma_j)z_I}\ \tilde{\Omega}_{mj}, \tag{2.101}$$

where $\tilde{\Omega} = \tilde{\mathbf{u}}(\tilde{\Gamma} + i\tilde{\mathbf{u}})^{-1}\tilde{\Gamma}$ with $\tilde{\Gamma}_{ij} = \delta_{ij}\gamma_j$ and

$$u_{ij} = \beta \int dy \phi_i^*(y) \phi_j(y) \exp(-q|y - y_I|). \tag{2.102}$$

The conductance for an infinite constriction containing an elastic scatterer as described in Equation 2.100 is expressed in terms of these matrices as

$$G_\infty = \frac{2e^2}{h} \sum_j^{\epsilon_j < E_F} \{1 + 2\frac{\text{Im}\ (\tilde{\Omega})_{jj}}{\gamma_j} + \frac{\text{Re}\ (\tilde{\Omega}^\dagger \tilde{\Gamma}^{-1} \tilde{\Omega})_{jj}}{\gamma_j}\}. \tag{2.103}$$

It should be noted that the effect of the evanescent waves with $\epsilon_j > E_F$ is included in the above formalism of G_∞. This is provided through the intersubband coupling in $\tilde{\mathbf{u}}$ and yields novel effects which do not exist in strictly 1D systems.

Chapter 3

Theory of Tunneling in Laterally Confined Systems

3.1 Tunneling in the presence of Lateral Confinement

Theory of tunneling was to a large extent formulated for laterally uniform systems. Thus, the lateral momentum of electrons tunneling through a barrier is conserved and the 1D Schrödinger equation is solved for the longitudinal energy

$$E_z = E - \frac{\hbar^2 |\mathbf{k}_{||}|^2}{2m}, \tag{3.1}$$

being the eigenvalue. Clearly, the transverse momentum $\mathbf{k}_{||}$ is associated with an effective potential in the Schrödinger equation. This is very natural, since only the momentum along the direction of tunneling barrier is involved in the tunneling process. Although for STM the tunneling barrier is not laterally uniform, still by using Bardeen's transfer Hamiltonian method[68] it is possible to have a similar result, and as described in Section 1.3 this is done by Tersoff and Hamann.[67] Their tunneling current expression is appealingly simple and it is possible to interpret a wide spectrum of STM results by using Equation 1.31. The local nature of STM probing is manifested in this expression. The confinement of the

current carrying states, however, is not explicitly taken into account. Varying the radius of the tip or the distance between the tip and sample one gets the usual dependence of tunneling current on these parameters[67] which is familiar from 1D tunneling. We believe that the lateral confinement in the tunneling barrier has to be an integral part of the theory of STM. This is not solely a belief, and there are evidences that such a formalism explains some novel phenomena taking place better than the conventional Tersoff-Hamann result, as explained in the following sections.

The most important drawback of Tersoff-Hamann theory[67] of STM, is the neglection of the tip-sample interaction. As mentioned in Section 1.3, the wave functions of the bare sample and tip suffice to calculate the tunneling current in this context. Earlier, Tekman[127] and Tekman and Ciraci[128,129] showed that at small tip-sample distance the interaction between the tip and sample leads to formation of tip induced localized (or resonance) states (TILS) which in turn have important contributions to the tunneling current. This is also apparent from the fact that for small tip-sample distance an adhesive force[130] attracts the electrodes to each other which can be interpreted in terms of a gradual collapse of the tunneling barrier. Thus, a correct theory of STM has to include the independent electrode regime (where the Tersoff-Hamann result[67] applies), the TILS regime and the electrical contact regime (for which the potential barrier is totally collapsed), as well. For an atomic scale theory, the mechanical contact of the electrodes[69] is another regime that has to be considered in addition to these.

As the tip approaches the sample, not only the magnitude but the real-space extent of the tip-sample interaction varies. This can be verified even by considering the Tersoff-Hamann result.[67] For large tip-sample distance the tunneling current is only weakly confined and the corrugation is small. Decreasing the tip-sample distance, the tunneling current gets confined in a very narrow region of space which yields atomic corrugation. This is, however, only a geometric effect in the context of Tersoff-Hamann theory. The collapse of the potential barrier has a similar trend and the physical picture changes, as the tip-sample distance decreases, from a spherical tip on a sample surface to a strongly

confined channel between the tip and sample. Using this point of view it is possible to form an analogy between the STM operating in the TILS regime and a pinched off QPC through which electrons can tunnel. In what follows we reformulate the theory of ballistic transport in a QPC, as explained in Section 2.1, in such a way that it is applicable to STM for small tip-sample distance.

We describe the metal electrodes by use of the jellium approximation, which is reliable if the tip and sample have simple Fermi surfaces without complicated features. That is, this method is not applicable *per se* to semimetal and semiconductor samples. Note that, STM is an intrinsically 3D system, as compared to 2D QPC studied in Chapter 2. Thus the potential energy $V(\mathbf{r})$ in the vicinity of the tip is given by

$$V(\mathbf{r};d) = V_s(z;d) + V_c(\mathbf{r}_{||},z;d), \tag{3.2}$$

where $\mathbf{r}_{||}$ is the lateral projection of the position vector $\mathbf{r}$, d is the tip-sample distance and the dependence of the potential on the interelectrode separation is explicitly taken into account. This is, in fact, another difference between the QPC and STM. For STM the system is defined externally by the tip-sample distance d. Recalling the discussion at the beginning of Section 2.2, one can argue that carrying out a parametrical analysis for QPC is sufficient to understand the basic ingredients of the system. However, for STM the quasi-1D nature of the structure is not reflected in the experimental results explicitly, as compared to the quantization of the conductance in the QPC, so that we have to calculate the experimentally relevant quantities.

The longitudinal and confining parts of the potential in Equation 3.2 are similar to those given in Equation 2.1. For the system at hand it is not possible to define sharp edges for the orifice as it is done in Equation 2.1. Nevertheless, it is clear that the confining potential is appreciable only in the vacuum barrier and the longitudinal potential $V_s(z;d)$ is calculated by using the jellium approximation so that going into the electrode it approaches to zero. Note that, in the vicinity of the tip the states are confined in two dimensions (i.e., in the $\mathbf{r}_{||}$ plane). Thus, the lateral momentum has a different quantization scheme as compared to that

for QPC. For the time being we assume that two quantum numbers forming a vector n characterizing the subband. The corresponding subband wave functions are $\xi_{\mathbf{n}E}(\mathbf{r})$ and $\bar{\xi}_{\mathbf{n}E}(\mathbf{r})$ for right- and left-going probability current densities, respectively. The current carrying solution $\psi_{\mathbf{k}}(\mathbf{r})$ of the Schrödinger equation for incidence from left (tip) with wave vector $\mathbf{k} = k_z(\mathbf{k}_{||})\hat{z} + \mathbf{k}_{||}$ is given by

$$\begin{aligned}
\psi_{\mathbf{k}}(\mathbf{r}) &= e^{ikz}\, e^{i\mathbf{k}_{||}\cdot\mathbf{r}_{||}} + \int d\mathbf{k}'_{||}\, e^{-ik_z(\mathbf{k}'_{||})z}\, e^{i\mathbf{k}'_{||}\cdot\mathbf{r}_{||}}\, A_{\mathbf{k}}(\mathbf{k}'_{||}), && z < -Z_t, \\
&= \sum_{\mathbf{n}} \{\xi_{\mathbf{n}E}(\mathbf{r})\Theta_{\mathbf{nk}} + \bar{\xi}_{\mathbf{n}E}(\mathbf{r})\Delta_{\mathbf{nk}}\}, && -Z_t < z < Z_s, \\
&= \int d\mathbf{k}'_{||}\, e^{ik_z(\mathbf{k}'_{||})z}\, e^{i\mathbf{k}'_{||}\cdot\mathbf{r}_{||}}\, B_{\mathbf{k}}(\mathbf{k}'_{||}), && Z_s < z, \qquad (3.3)
\end{aligned}$$

where $-Z_t$ and Z_s are the positions of the edges of the constriction at the tip and sample sides, respectively. They are defined somewhat arbitrarily, since the only restriction is that the potential in the semiinfinite half-spaces $z < -Z_t$ and $z > Z_s$ has to be zero. The longitudinal momentum is $k_z(\mathbf{k}_{||}) = \sqrt{2mE/\hbar^2 - |\mathbf{k}_{||}|^2}$ with positive imaginary part, as it is defined just after Equation 2.4. For this wave function the solution procedure is the same as that for QPC, as described in Section 2.1.

It is clear that for the problem at hand it is not possible to talk about a uniform constriction in the sense that is defined at the beginning of Section 2.2. The reason for this is the presence of the longitudinal potential $V_s(z)$ which is the tunneling barrier for the laterally uniform system. That is, it is necessary to use the transfer matrix method described in Section 2.3 in numerical calculations. Consequently, the tunneling conductance can be written in conjunction with Equation 2.18 as

$$G_t = \frac{e^2}{2\pi^2 h} \int_{F.S.} \frac{d\mathbf{k}_{||}}{k_z(\mathbf{k}_{||})} \{\tilde{\Theta}^{\dagger}_{\mathbf{k}} \tilde{\Gamma}_R \tilde{\Theta}_{\mathbf{k}} - \tilde{\Delta}^{\dagger}_{\mathbf{k}} \tilde{\Gamma}_R \tilde{\Delta}_{\mathbf{k}} + 2\mathrm{Im}[\tilde{\Theta}^{\dagger}_{\mathbf{k}} \tilde{\Gamma}_I \tilde{\Delta}_{\mathbf{k}}]\}, \qquad (3.4)$$

where the integral is carried over the Fermi hemisphere with positive k_z and all the matrices have their usual meanings. The coefficients $\tilde{\Theta}$ and $\tilde{\Delta}$ can be calculated in any segment used for the transfer matrix, since the current is conserved. We want to stress once more that the expression in Equation 3.4 is for infinitesimal bias. Note that, this expression for the tunneling conductance applies equally well to the point contact regime, since the propagating states are included in the

first two terms in the brackets. The tunneling contribution is the third term and is important for the independent electrode and TILS regimes of STM.

At this point it is in order to compare the present approach with the earlier studies on STM. It is known that the jellium approximation is appropriate to calculate the tunneling current as a function of d, but does not convey any information regarding the lateral confinement of the current. In order to resolve interactions on the atomic scale an individual atom was attached on one of the jellium surfaces.[131-133] Even this approach provides limited applicability in the analysis of the STM. On the other hand, by using the self-consistent field (SCF) calculations for a periodically repeating tip-sample system, the tip-sample interactions can be resolved[134-136] on the atomic scale and information about the charge density and potential at the sample surface can be obtained. In this case the calculation of the tunneling current is, however, hindered, since the size of the supercell representing the repeating tip-sample system is finite and thus states in the k-space are discretized. The formalism developed above, which is a combination of these two methods, is both computationally simpler than the SCF calculations and superior to the simple jellium approximation. The use of the jellium approximation is apparent in Equations 3.3 and 3.4. The SCF calculations, in turn, are used to find the form and parameters of the potential $V(\mathbf{r}_{||};d)$ in the vacuum region. This way the present method becomes simple (due to the use of jellium approximation) and realistic (due to the use of the SCF results).

The details of the SCF calculations for the Al tip-Al sample system are given in References 134–136, and the results are summarized in Appendix 3.A. The main conclusion of these studies, as long as the application of the present formalism is concerned, is that simulating the longitudinal potential by the bimetallic-junction potential in the jellium approximation is plausible and the confining potential resembles to a 2D parabola. Although the discrete nature of the atoms used in the SCF calculations yield some deviations, our choice for the model potential in the vacuum region is

$$V(\mathbf{r};d) = \phi_m(z;d) + \alpha(z;d)\rho^2\theta(z+Z_t)\theta(Z_s-z) \qquad (3.5)$$

with $\rho = |\mathbf{r}_{||}|$. Here $\phi_m(z;d)$ is the bimetallic-junction potential for two jellia,[137] edges of which are at $z = 0$ (tip) and $z = d$ (sample), and $\alpha(z;d)$ is found by fitting a parabola to the self-consistent potential. Since the confining potential is isotropic in $\mathbf{r}_{||}$-plane, the lateral subband wave functions are products of two harmonic oscillator solutions

$$\phi_{\mathbf{n}}(\mathbf{r}_{||}) = \phi_{n_x}(x)\phi_{n_y}(y), \tag{3.6}$$

with $\mathbf{n} = (n_x, n_y)$ and $\epsilon_{\mathbf{n}} = \epsilon_{n_x} + \epsilon_{n_y} = \hbar\omega(n_x + n_y + 1)$, $\omega^2 = 2\alpha/m$. In the above, the entities with vector and scalar subscripts are 2D and 1D quantities, respectively. However, from now on we will denote the 2D subbands by using a single index $n = n_x + n_y$ for the sake of simplicity. An important quantity that we refer frequently in the following sections is the lateral width of the wave functions $w = (\hbar^2/2m\alpha)^{1/4} = (\hbar/m\omega)^{1/2}$. In the following sections we use the potential given in Equation 3.5 together with the results of the SCF solutions for the Al tip-Al sample system to investigate some novel phenomena taking place in STM.

3.2 Transition to Point Contact in Scanning Tunneling Microscopy

Transition to point contact in STM was first studied by Gimzewski and Möller.[69,70] Two striking features of their results are that the tunneling current saturates as the tip is brought closer to the sample and it increases discontinuously at a certain tip-sample distance. The analysis of this discontinuous jump showed that it is due to the adhesion of tip and sample, which happens when the tip-sample system becomes unstable and the tip elongates towards the sample to form a mechanical contact. This is verified by the hysteric behavior of the tunneling current after this discontinuous jump and poor reproducibility of the jump. Interestingly, some oscillations have been observed[70] in this hysteric region as the tip-sample distance (actually the voltage applied to the z-piezo) is varied.

Gimzewski and Möller[69] gave an estimate for the radius of the contact area, which lies in the range of λ_F. If so, the observed transport beyond the discontinuity has to be associated with the ballistic transport in the quantum regime. García[138] was the first who pointed out that the point contact in STM is related to the ballistic transport of electrons through a QPC. He attributed the oscillations above the discontinuous jump to occupation of subbands in the constriction formed at the point contact. Later, two studies on the transition to the point contact were reported[132,139] which uses the solution of Schrödinger equation for model tip-sample systems. Lang[132] simulated the point contact by two jellium electrodes, one of them having an adsorbed Na atom on it and thus representing a single atom tip. He found that the conductance G_t saturates at the value $\eta 2e^2/h$ and forms a plateau when d is in the range of twice the distance from the Na core to the jellium edge of the tip electrode. The value of η is only 0.4 for Na, and is found to depend on the identity of the tip. Within a tight-binding approximation and using the nonequilibrium Green's function method Ferrer and coworkers[139] also found that G_t saturates at $\lesssim 2e^2/h$. Moreover, in both calculations,[132,139] the mechanism related to the experimentally observed increase of G_t following a discontinuity was not included.

Ciraci and Tekman[140] used the present formalism together with the SCF results obtained for Al tip-Al sample system to understand the mechanism responsible for the transition from tunneling to point contact and the reason for oscillations observed after the discontinuous jump. In this section we follow the same line of reasoning to analyze the problem in more detail. The experiment[69] was carried out by using an Ir tip and a polycrystaline Ag sample. In a very recent experiment[141] both the tip and the sample are taken to be made of Ir. In order to prevent the possible complications arising from the presence of different materials and d-type electronic states, we use a Ag tip-Ag sample system using the jellium parameters given by Smith.[142] Whereas the SCF results[135,136] are for Al tip and sample. Note that, one of the most important parameters of the system is obtained by making use of the SCF results, namely $\alpha(z;d)$. As a first approximation we use a constant $\alpha(d)$ throughout the constriction lying between

$-Z_t = -a_o/2$ and $Z_s = d$, where a_o is the interlayer distance for Ag. This choice of parameters simplify the calculations to a great extent and, in addition, present a quite fair approximation to the realistic case. The confinement parameter $\alpha(d)$ is found by averaging $w(d)$ found for the ontop- (T) and hollow-site (H-site) positions of the tip for the Al sample and scaling it with the ratio of the atomic radii of Ag and Al as defined by Harrison[143] $r_a = (3Z/4\pi^2 n)$, where Z is the valency and n is the density of electrons. Although Ag and Al have one and three valence electrons, respectively, the corresponding atomic radii are almost the same. Calculating $w(d)$ for distances used in the SCF calculations, we fitted these discrete values to a line and thus obtained the lateral confinement for an arbitrary tip-sample distance. To this end, the confinement parameter w is given by

$$w(\mathring{A}) = 1.1361 + 0.3760 \times d(\mathring{A}), \tag{3.7}$$

and the subband energies can be found by using $\hbar\omega(\text{eV}) = 3.82/[w(\mathring{A})]^2$. The longitudinal potential $\phi_m(z;d)$, on the other hand, is calculated by using the approximation scheme described by Ferrante and Smith.[137] This completes the determination of the potential in Equation 3.5. Next, we calculate the tunneling conductance G_t as given in Equation 3.4.

The results of our calculations are presented in Figure 3.1. For large d, the effect of the lateral confinement is small and the STM operates in the independent electrode regime. Therefore, the tunneling conductance changes exponentially with distance. For intermediate values of distance, however, the exponential dependence is no more valid. The reason for this, actually, is twofold. It is clear that the maximum transmission probability through the barrier is unity, thus the conductance can not increase indefinitely in an exponential form. More importantly the effective barrier height[‡] $\phi_{eff}(d) = \max\{\phi_m(z;d) - E_F\} + \epsilon_0(d)$ has an increasing contribution due to the stronger lateral confinement for smaller tip-sample distance. In short, the current becomes smaller than that would be

[‡]For STM problems we use a single effective barrier, namely that for the lowest subband, as long as there is not a reason for not doing so. This is for simplifying the comparison with the existing literature, since there is a well established use of the term "effective barrier" in works related to the STM.

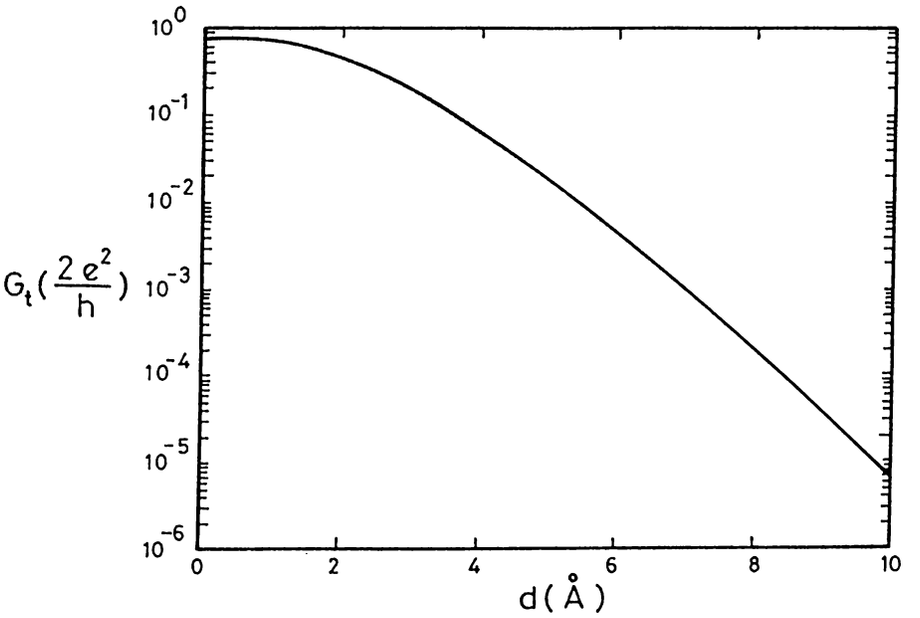

Figure 3.1: Tunneling conductance for scanning tunneling microscopy including the lateral confinement effects
The conductance as a function of the distance between the jellium edges. The lateral confinement parameter is given by Equation 3.7.

expected for the exponential dependence. Finally, for $d \simeq 1$ $\AA$ the tunneling conductance saturates for $G_t \sim 0.75$ $(2e^2/h)$, which corresponds to a resistance $R_t \sim 17$ KΩ. These results are consistent with those obtained by other studies,[132,139] namely the conductance associated with a uniform orifice set up by a single atom at the vertex of the tip has a value less than $2e^2/h$. Lang[132] attributed this to chemical identity of the foremost atom at the apex of the tip. Ferrer and coworkers,[139] on the other hand, find almost one quantum of conductance for the maximum value of G_t, due to the perfect matching between the tip and sample. This agreement between three calculations employing completely different methods is striking and give evidence to the importance of the lateral confinement effects in STM at small tip-sample distance.

Next we compare the results presented in Figure 3.1 with those obtained in experiments.[69,141] As depicted above, the main feature of the experimental

results is the discontinuous jump in tunneling conductance. The jump occurs for $R_t \sim 40$ KΩ and after the jump the resistance drops to $R_t \sim 10 - 18$ KΩ which is of the order of the quantum of conductance. Accordingly, the saturation shown in Figure 3.1 can not be associated with the saturation observed in the experiments. As evidenced by the hysteric nature of current[69,70] and force-gradient measurements,[141] the discontinuous jump is related with the adhesion of the tip to the sample. Thus, the present method can not be used to analyze the nature of this discontinuous jump. Nevertheless, we used different confinement parameters and tip sizes in order to analyze the how the transition from tunneling to point contact takes place and whether a plateau appears just below the discontinuity. We find that decreasing the effect of the lateral confinement (i.e., increasing w) the saturation shifts to larger d. On the other hand, increasing (decreasing) the length of the apex of the tip (i.e., Z_t) the saturation completely disappears and as $d \to 0$ the current begins to decrease (increase) beyond the saturation value. Therefore, we conclude that approach to the point contact sensitively depends on the tip geometry and may not be easily reproducible.

In order to compare the present results with the experiment in detail we calculate the apparent barrier height

$$\phi_A = \frac{\hbar^2}{8m} \left(\frac{\partial \ln I}{\partial d} \right)^2 , \tag{3.8}$$

which is equal to the real barrier height $\phi(d) = \max\{\phi_m(z;d) - E_F\}$ for planar electrodes separated by a wide potential barrier. The result is given in Figure 3.2. Clearly, the apparent barrier height decreases with decreasing d due to the above mentioned saturation effects. Lang[133] calculated the apparent barrier height for a single atomic junction and had a qualitatively similar result. The only discrepancy is that according to Lang the apparent barrier height drops rapidly from a value close to or above the work function of the metal $\phi_A \sim \Phi$, to zero. We find that the drop is rather gradual and at the position of the current discontinuity it is ~ 0.5 eV, which is close to the measured value[69] of ~ 1 eV. In the independent electrode regime, however, it is close to the work function of Ag, $\Phi = 3.2$ eV. In the same figure the real barrier height $\phi(d)$ and the apparent barrier height

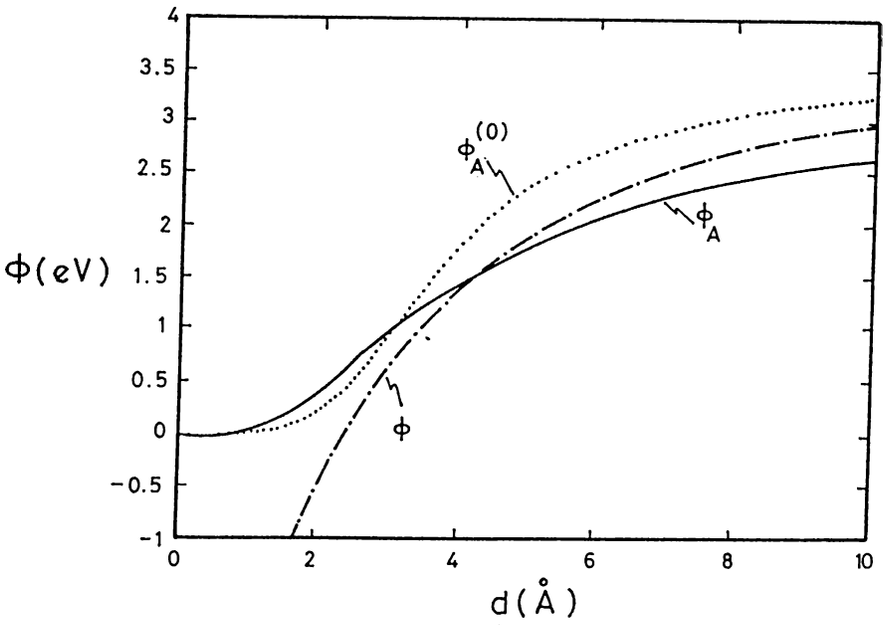

Figure 3.2: Apparent and real barrier heights for scanning tunneling microscopy The full, dash-dotted and dotted curves show the apparent potential calculated from Figure 3.1, the saddle point value of the real potential and the apparent potential for flat jellium surfaces, respectively.

in the absence of the lateral confinement $\phi_A^{(0)}(d)$ are also shown for comparison. Clearly, $\phi_A(d)$ is smaller than $\phi(d)$, and consequently than $\phi_{eff}(d)$, for the TILS regime of STM, i.e., $d \gtrsim 4\mathring{A}$. This is in contrast to what is found for a strictly 1D system.[132] The reason for this is the deviation from the exponential dependence as described above. Note that, the apparent barrier height is not related to any kind of real potential barrier if the electrodes are close to each other or the real (or effective) potential barrier is low, but only constitutes an alternative way of representing the current. Nevertheless, it contains additional information as compared to a single $I - d$ graph and it is desirable to have the results for ϕ_A together with the tunneling current.

According to the experiment[69] it appears that the ballistic regime starts subsequent to the structural instability coexisting with the discontinuity in the current. Under normal circumstances, if one continues to push the tip further, the

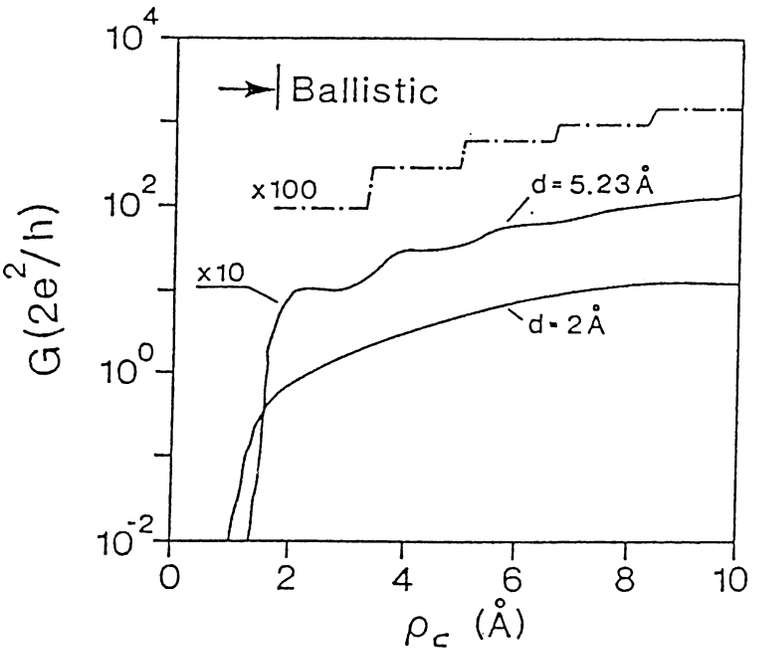

Figure 3.3: Conductance of a point contact in scanning tunneling microscopy The conductance as a function of the radius of the point contact, the length of the constriction being a parameter. The dotted line shows the ideal quantization of conductance. The curves are offset for clarity.

distance between the outermost tip and sample layers do not change appreciably, but the contact aperture expands with an enhanced plastic deformation followed by the adhesion of several atoms. The actual form and size of the contact after the point of mechanical instability is uncertain and depends on several parameters (such as the detailed atomic structure at the apex of the tip, intra and interelectrode interaction energies). Nevertheless, the variation of conductance as a function of the tip displacement can be related to the radius of the orifice, which normally increases with the continuing plastic deformation. Depending on the aperture of the orifice several subband states can be occupied. Each occupied subband contributes to G_t by $2e^2n/h$ (n is the degeneracy). The perfect quantization with sharp steps occurs if the orifice is longer than λ_F, however. In what follows we extend our model to include the plastic deformation region, where a significant hysteresis is observed in the excursion of the tip.[70] We calculate

the conductance for $\phi_m(z;d) = 0$ and as a function of the radius of the aperture $\rho_c = \sqrt{E_F/\alpha}$. The results are presented in Figure 3.3. Expectedly, for long orifices (i.e., $d \gtrsim \lambda_F$) the conductance is approximately quantized and well defined steps and plateaus are apparent in G_t versus ρ_c curve. On the contrary, for short orifices (i.e., $d \lesssim \lambda_F/2$) the tunneling effects are important and the steps are smeared out yielding the quantum Sharvin conductance. For this case the conductance curve is featureless and increases uniformly. Recalling that the interlayer distance for Ag(111) is $a_o = 2.36$ $\AA$ and the Fermi wavelength is $\lambda_F = 5.23$ $\AA$, for the system at hand it is not possible to observe quantization of conductance after the point contact is initiated. The oscillations seen in some of the experiments,[70] therefore, can not be associated with a quantum interference effect. The reason for these oscillations is most probably is the nonuniform enlargement of the contact are due to the irreversible displacement of atomic cores.

Finally we wish to point out one important observation about using the jellium parameter $r_s = (3/4\pi^2 n)$, which is the radius of the sphere containing a single electron, as the scaling parameter instead of the atomic radius r_a. The electron densities of Al and Ag are quite different and for this case the scaling have an important effect on the results by decreasing the width w nearly by 50%. The conductance calculated thereof, in turn, would be very small compared to that given in Figure 3.1 due to the large ϕ_{eff}. Therefore, one can conclude that this scaling is not a proper one as far as the consistency with the experiment is concerned. Thus, the atomic nature of the small tip-sample distance STM is manifested in an indirect yet rather striking way.

3.3 Atomic Corrugation on Al(111) Surface

Another interesting result obtained by using STM is the atomic corrugation observed for nominally flat metal surfaces.[71,72] The (111) surfaces of noble and simple metals are usually have very small charge and potential corrugation as a result of their close packing. Recent STM experiments,[71,72] however, yield height corrugations as large as 0.3 $\AA$. This is much larger than that one can deduce from

the charge density of the free surfaces, which is the theoretical result for Tersoff-Hamann theory of STM.[67] The anomalous corrugation for the Au surface was attributed to a surface state lying 0.5 eV below the Fermi level.[71] The electronic structure of Al, on the other hand, is not suitable to relate any special state to the observed anomalous corrugation. Winterlin and coworkers[72] argued that the STM corrugation of the Al(111) surface is enhanced by the elastic deformation of the tip, which is induced by the attractive forces between two electrodes. Recent theoretical calculations by Ciraci and coworkers,[135] however, are at variance with this interpretation. Their SCF results showed that the observed corrugation is reduced, but not enhanced, by the tip induced elastic deformation. They also showed that due to the tip-sample interaction the electronic structure of the electrodes change drastically and depending on the lateral position of the tip. For the reasons pointed out above, however, they were not able to calculate the tunneling current using these SCF results. Two recent explanations for the anomalous corrugation,[144,145] on the other hand, did not address the question of the tip-sample interaction in spite of the fact that the interelectrode distance is small.

In this section we present an explanation for the anomalous corrugation of Al(111) surface obtained by STM[146] using the formalism developed in Sections 2.1 and 3.1. That is, the origin of anomalous corrugation for close-packed metal surfaces is sought in the tip-sample interaction and the resulting lateral confinement of the current carrying states. First we analyze the results of the SCF calculations summarized in Appendix 3.A and determine the model potential in Equation 3.5. In Figures 3.4 and 3.5 the self-consistent potential $V(\mathbf{r};d)$ along $z=0$ line are shown for T and H positions of the tip, respectively. The most important observation is the larger barrier width for the hollow-site position of the tip as compared to the ontop-site position, which is a result of the position dependent tip-sample interaction. We quantify this difference between the barrier widths in terms of the quantity $\Delta\xi(E;d) = \xi_H(E;d) - \xi_T(E;d)$, where the individual barrier widths for a given energy are defined as $\xi_H(E;d)$ and $\xi_T(E;d)$ for the hollow- and ontop-sites, respectively. Note that, $\Delta\xi(E;d)$

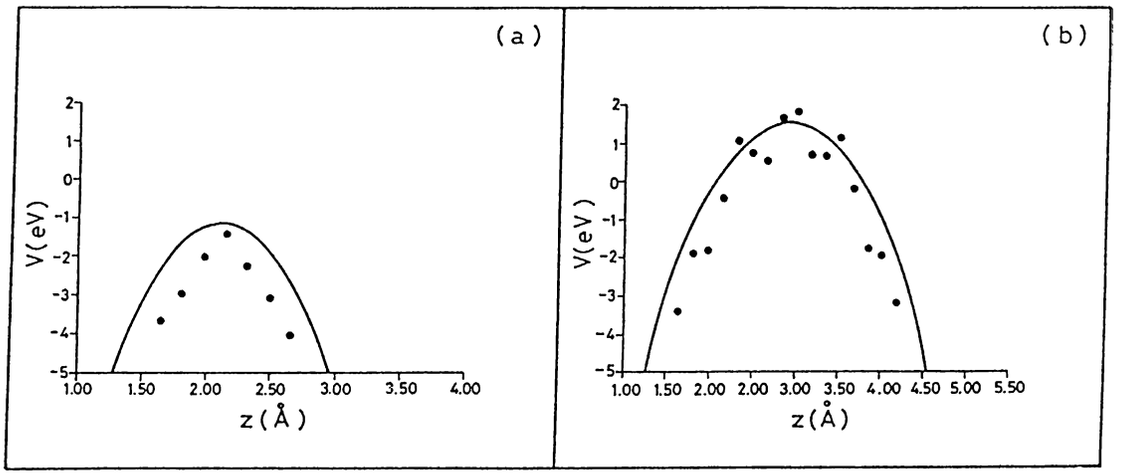

Figure 3.4: Self-consistent potential along the longitudinal direction: Ontop-site
The circles represent the SCF results for the potential. The tip-sample interatomic distance is (a) 8 and (b) 11 a.u. The lines represent the potential profile used in the calculation of conductance.

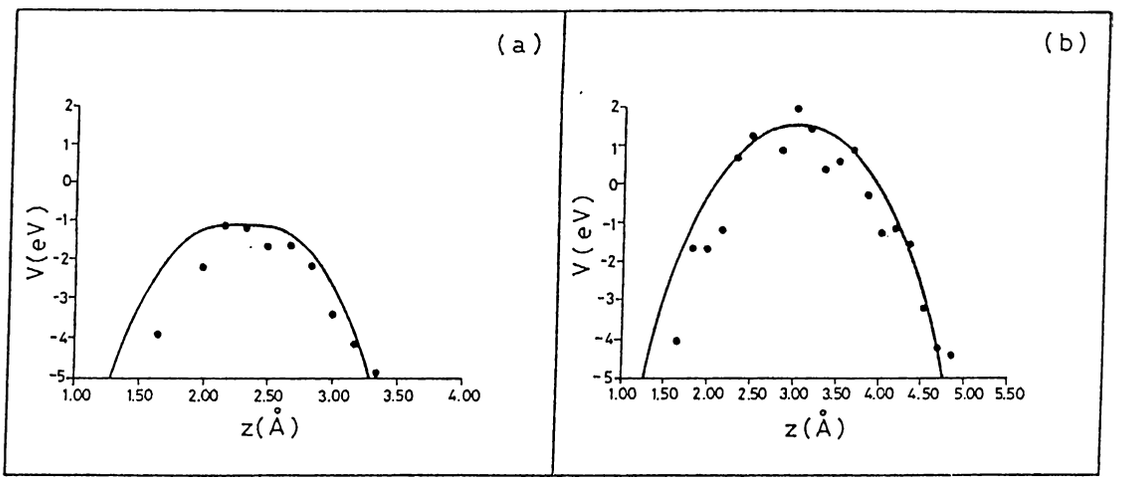

Figure 3.5: Self-consistent potential along the longitudinal direction: Hollow-site
The circles represent the SCF results for the potential. The tip-sample interatomic distance is (a) 8 and (b) 11 a.u. The lines represent the potential profile used in the calculation of conductance.

can be thought as a potential corrugation induced by the tip-sample interaction.

In contrast, the maximum of the potential barrier is not sensitive to the lateral position of the tip to an appreciable extent.

The effect that enhances $\Delta\xi$ at small d can be found in the tip-sample interaction. The surface potential of the free sample $V_s(\mathbf{r})$ is dominated by the exchange-correlation potential (which in the local density approximation is proportional to $\rho_s^{1/3}$) and the potential corrugation $\Delta\xi_s(E;\infty)$ is still a small quantity ($\ll 0.1$ Å). Apparently the measure obtained by He-scattering experiments is $\Delta\xi_s(E_{\mathrm{He}};\infty)$, E_{He} denoting the energy of the He atom. Only very close to the surface or for energies far above the Fermi level, due to the coulombic potential (i.e., attractive core and repulsive Hartree potential) the corrugation $\Delta\xi_s(E;\infty)$ is comparatively larger. Nevertheless, these conditions are not accessible neither with He-scattering nor with STM operating in the independent (noninteracting) electrode regime. As pointed earlier by Tekman and Ciraci,[128,129] the tip and sample states are combined to yield tip induced localized states in STM for small d. This induces substantial local modifications in the charge distribution between the two electrodes and this redistribution of charge is strongly site-dependent[135,136] for an atomically sharp tip. This site-specific rearrangement of the charge at small d amplifies the corrugation of the charge density $\Delta\rho(d)$, and thus leads to a large value for $\Delta\xi(E;d)$.

In the model potential Equation 3.5 the bimetallic junction potential $\phi_m(z;d)$ as described by Ferrante and Smith[137] for the T-site[‡]. Note that, we use Al jellium parameters for both the tip and sample. This is consistent with the experimental arrangement,[71,72] however. In experiments atomic corrugation was observed only after reshaping the tip by applying a large bias and thus having a material transfer between the electrodes, most probably Au or Al transfer to the apex of the tip.[72,135] The potential for the H-site, on the other hand is calculated by using the same jellium potential, with an elongation at the saddle point of length $\Delta\xi(E_F-2\text{ eV};d)$ in order to take care of the tip-sample interaction effects. The potential corrugation induced by the tip-sample interaction is shown

[‡]We fit the potential obtained from SCF calculations to the jellium results so that the jellium parameters for Al is different than those given in Reference 142. We use $n_+ = 26.9\times10^{-3}$ a.u.$^{-3}$, $\beta = 1.122$ and $\Phi = 4.983$ eV.

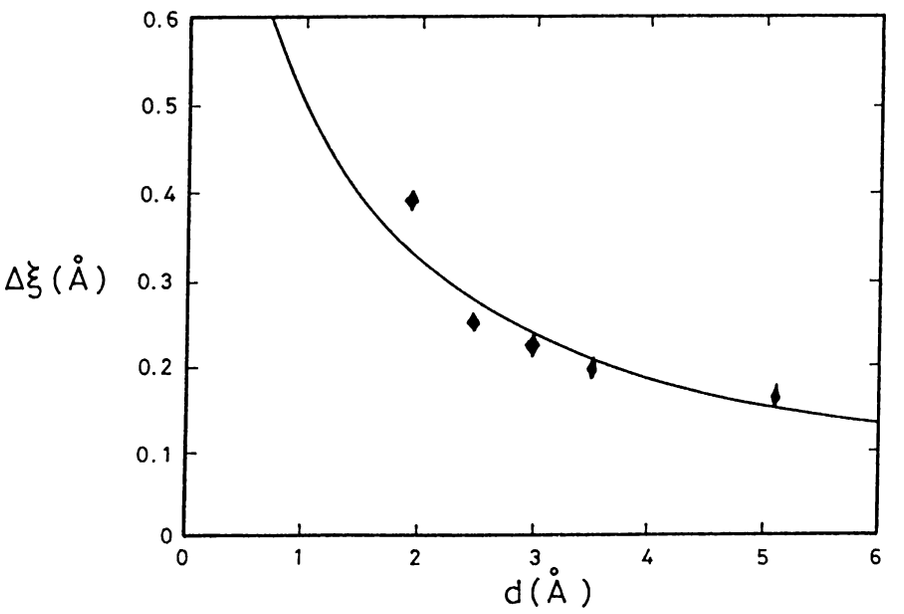

Figure 3.6: Potential corrugation induced by the tip-sample interaction
The potential corrugation as a function of the distance between the jellium edges. The diamonds denote the values found from SCF calculations and full line is the fit given by Equation 3.9.

in Figure 3.6. It is possible to fit a smooth function to the results obtained from SCF calculations and we find

$$\Delta\xi(\mathring{A}) = [1.0811 \times d(\mathring{A}) + 0.9171]^{-1}, \tag{3.9}$$

which is calculated 2 eV below the Fermi level. This choice of energy is not essential and the potential corrugation is constant for a wide range of energy below the Fermi level.

Next, we analyze the lateral confinement of the current carrying states. In Figures 3.7 and 3.8 the SCF potential in the transverse plane bisecting the line segment between the outermost tip atom and surface plane are shown for ontop- and hollow-site position of the tip, respectively. Apparently, the confinement gets stronger as the tip gets closer to the sample. In addition, the confinement for the T-site is stronger than that of the H-site. The structure superimposed on the

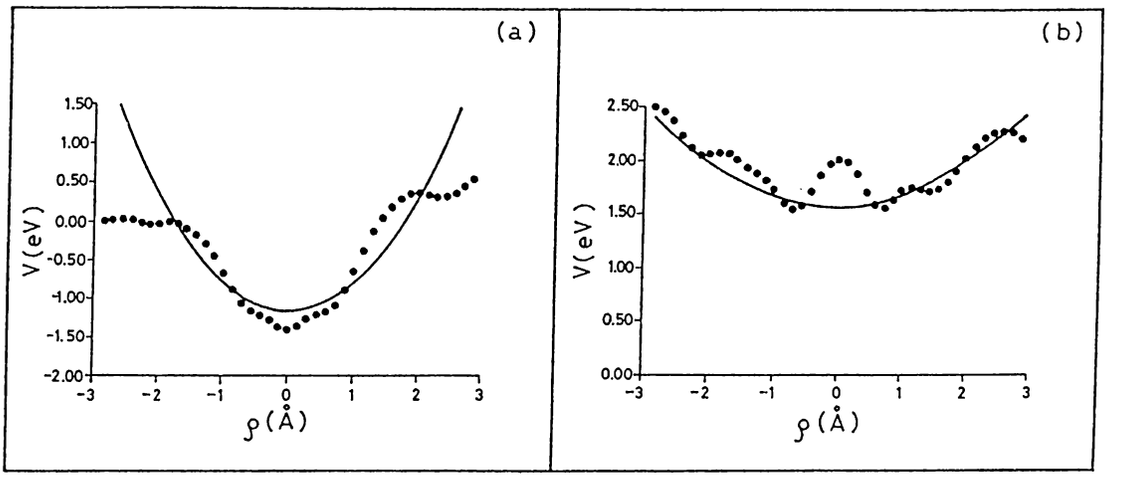

Figure 3.7: Self-consistent potential in the transverse plane: Ontop site
The circles represent the SCF results for the potential. The tip-sample interatomic distance is (a) 8 and (b) 11 a.u. The lines represent the potential profile used in the calculation of conductance

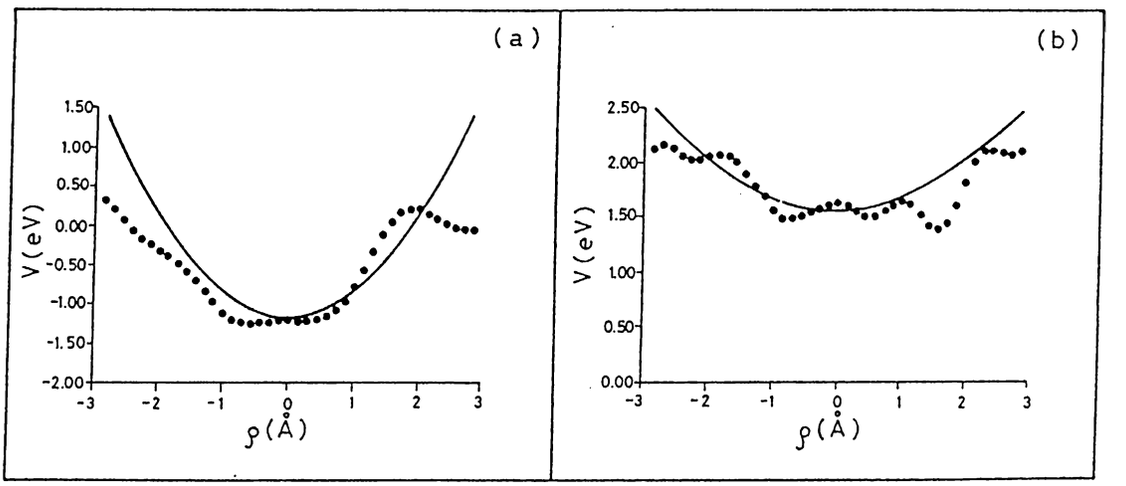

Figure 3.8: Self-consistent potential in the transverse plane: Hollow site
The circles represent the SCF results for the potential. The tip-sample interatomic distance is (a) 8 and (b) 11 a.u. The lines represent the potential profile used in the calculation of conductance

smooth background is caused by the discrete nature of the atoms. We fit these potentials to a parabola to find $\alpha(d)$ at the bisecting plane for T- and H-site. In

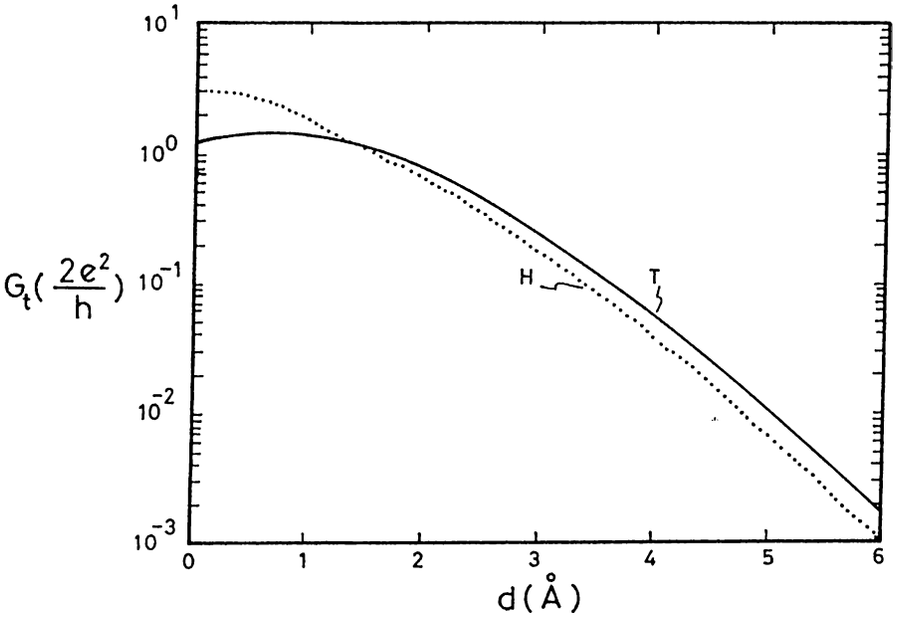

Figure 3.9: Tunneling current for Al tip-Al sample system
The conductance as a function of the distance between the jellium edges. The full and dotted lines denote the conductance obtained at the ontop- and hollow-site positions of the tip, respectively.

turn, the confinement parameter w is found as

$$w_T(\mathring{A}) = 0.9645 + 0.4270 \times d(\mathring{A}), \tag{3.10}$$

$$w_H(\mathring{A}) = 1.2051 + 0.3441 \times d(\mathring{A}). \tag{3.11}$$

We use these confinement parameters throughout the orifice extending from $-Z_t = -a_o/2$ to $Z_s = d$, which constitutes a good approximation to the real potential. This completes the determination of the parameters of the calculation of tunneling current.

Next, we calculate the tunneling conductance for the ontop- and hollow-site position of the tip by using Equation 3.4. Our results are presented in Figure 3.9. The curves resemble to that given in Figure 3.1 for transition to point contact, as expected. For large d ($\gtrsim 5$ Å), $\log G_t$ is approximately a linear function of distance. This indicates that the transport occurs via tunneling, and specifically

STM is in either TILS or independent electrode regime. In this range of d, the current at the T-site is larger than that at the H-site and yields corrugation of ~ 0.3 $\AA$. This value is in agreement with the experimental observation, since the calculated tunneling current is ~10-20 nA for $d \sim 5.5$ $\AA$ and for the bias voltage of 50 mV which are typical for the observed anomalous corrugation.[72] Note that, for increasing d the corrugation remains approximately constant. This is due to insufficient fit of $w(d)$ to the self-consistent results. For a more realistic form of w, the effective barrier at the T- and H-sites should merge into a single one and thus the corrugation would vanish. In the intermediate region $2 \lesssim d \lesssim 4$ $\AA$, the effect of increasing lateral confinement at the T-site becomes superior to that of increasing $\Delta\xi$ at the H-site. Hence, the measured corrugation decreases with decreasing d. However, in this region it is possible to have some instabilities as a result of the tip-sample interaction and it is hard to observe this decreasing corrugation experimentally. Finally, as shown in Figure 3.9, for $d \lesssim 2$ $\AA$ the current at the H-site exceeds that at the T-site. This implies that the corrugation is inverted at small d before the mechanical contact, and thus the hollow-site (rather than the atomic sites) appears as a protrusion in the STM images obtained by the topographic mode. For this small d regime, however, the deformation of the tip and sample has to be considered to get a more realistic picture. In addition, observation of this effect may be hindered by the mechanic instability of the system.

3.4 Focusing of Electron Beams

3.4.1 Focused Field Emission of Electrons from a Point Source

In the preceding sections it is shown that the lateral confinement of the current carrying states affects the nature of electron transfer in STM to a great extent. Another example for such events is the focused field emission of electrons from a point source. Recently, Fink[74] achieved the fabrication of stable tips terminating

in a single atom providing a charged particle source of atomic dimensions. Low-energy electron beam with current ~ 10 μA is obtained thereof by using much lower voltages ($\sim 10^3$ V) than in conventional field emitters. Moreover, in spite of the finite transverse momentum of the incident electrons, the beam is well focused so that the angle of spread is measured to be $\sim 2^o$ at the screen.[74] Since then field emission of electrons have been the focus of attention from both experimental[147] and theoretical[148–151] points of view.

The theoretical studies by García and coworkers[148,150] and Lang and coworkers[149] on the field emission from an atomic size source arrived at different conclusions for the mechanism of collimation. In the former one, the point source was simulated by a quasi-1D constriction connected to a free electron reservoir, and the effect of the applied field was represented by a triangular potential barrier at the exit of this constriction. They found that the triangular shape of the potential barrier is responsible for the beam focusing. Later, a more elaborate study on the focused field emission of electrons[149] concluded that in the presence of an applied bias a channel with a horn-like potential profile opens in front of the apex of the tip. In this channel the single-particle wave functions evolve adiabatically[81] so that their transverse momentum decreases, without any scattering among subbands. Accordingly, focusing was attributed to the horn-like form of the channel. However, their self-consistent jellium calculations[149] alone did not provide evidence for the adiabaticity of propagation.

In what follows we use a model potential of the form given in Equation 3.5 in order to analyze the mechanism of focused electron emission from point sources.[152] Clearly, one has to include the effect of the electric field $\mathbf{F} = -F\hat{z}$ in the bimetallic junction potential $\phi_m(z)$. This is achieved by adopting the approximation method of Orosz and Balaźs,[153] namely by rigidly shifting the charge density of the electrons with respect to the jellium edges. The region for the lateral confinement, on the other hand, is different than those in Sections 3.2 and 3.3. In order to emphasize this difference we rewrite Equation 3.5 in a different way

$$V(\mathbf{r};F) = \phi_m(z;F) + \alpha(z;F)\rho^2\theta(z+r_t)\theta(d-z), \tag{3.12}$$

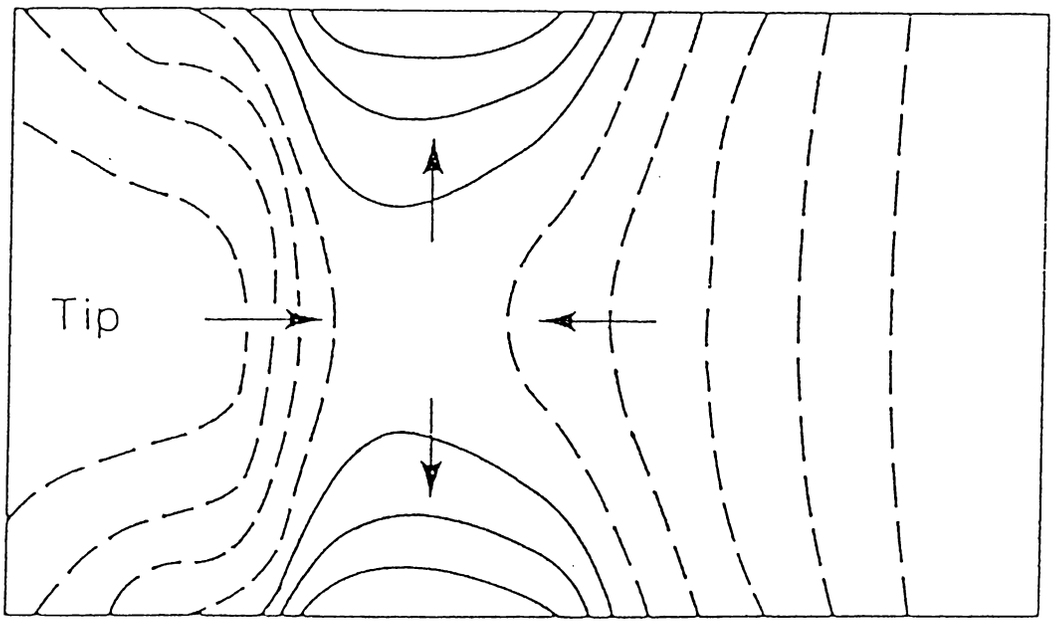

Figure 3.10: Potential profile for an electron emitting point source
The contours of constant potential are calculated for $L = 30$ Å, $d = 10$ Å, $r_t = r_h = 4$ Å (these parameters are used throughout this section). The electric field is 2 V/Å and the confinement parameters are 0.2, 0.5 and 0.02 ev/Å^2, at the base of the tip, around the saddle point and at the aperture to the vacuum, respectively, as described in the text. The contour plots are calculated in a plane $\rho \leq 4.25$ Å and $-4 \leq z \leq 10$ Å. The contour spacing is 2.5 eV and the dashed contours have potential below the Fermi level.

where $\phi_m(z;F)$ is calculated for two jellium edges located at $z = 0$ (emitter) and $z = L$ (screen), with a finite bias in between. The electric field asymptotically approaches to F away from the electrodes. The lateral confinement extends from $z = -r_t$ (base of the tip) to $z = d$ (aperture to the vacuum). $\alpha(z;F)$ is fixed to simulate the shape of the apex of the tip for $-r_t \lesssim z$ and the horn-like opening at the vacuum side $z \lesssim d$. The horn-like connections at the tip and vacuum sides are formed by varying the confinement parameter $w = (\hbar^2/2m\alpha)^{1/4}$ smoothly between the end points of the horns. Apparently, the determination of the potential from first principles, as it was done in Sections 3.2 and 3.3, is not an easy task due to the limitations of the local density approximation for low

electron densities. Therefore, we carry out a parametrical analysis aiming at a clear understanding of the mechanism of focusing. Clearly, the model potential given in Equation 3.12 is not identical to the actual one, but is appropriate to simulate a point source. The potential profile for a typical case is illustrated in Figure 3.10. Clearly, it resembles to those found by using self-consistent jellium calculations,[149] in addition our method is both transparent and computationally simple.

The current carrying states can be found by using the transfer matrix method described in Section 2.3 and the emission current is found by integrating the resulting energy current density $J(E;F)$, which is given by an expression similar to Equation 3.4,

$$I(F) = \int_0^{E_F} dE\ J(E;F), \tag{3.13}$$

where we take the lower limit for integration as the bottom of the conduction band of the tip electrode, since the bias applied between the emitter and screen is larger than E_F by at least an order of magnitude. An important problem is to quantify the focusing of an emitted electron beam. In the experiments[74,147] the emitter and screen are separated by macroscopic distances ($\lesssim$ 20 cm.) and the degree of focusing is related to the size of the spot on the screen. The propagation of the electron beam through this vacuum region can be analyzed quantum mechanically.[151] However, the subject matter of this study is the microscopical mechanism responsible for focusing. That is, the far field region is ignored and only the emitter and its vicinity is included in the model potential in Equation 3.12. We use the lateral momentum spectrum of the emitted beam as a measure for the collimation. To this end, we calculate the expected value of the lateral momentum squared just at the aperture of the horn, i.e., $z = d$. The wave function at the opening is given by[‡]

$$\psi_{\mathbf{k}_i}(\mathbf{r}) = \int d\mathbf{k}_{||}\ e^{ik_z(\mathbf{k}_{||})z}\ e^{i\mathbf{k}_{||}\cdot\mathbf{r}_{||}}\ B_{\mathbf{k}_i}(\mathbf{k}_{||}). \tag{3.14}$$

[‡]Here, we assume that the potential is flat for $z > d$ for the sake of computational simplicity. The same assumptions used in calculation of $I(F)$ as well. This approximation does not affect the results very much as far as it is considered in conjunction with the experimental configuration.

$B_{\mathbf{k}}$ can be calculated by using the appropriate forms of Equations 2.13 and 2.37. The expectation value of the lateral momentum squared for incident plane waves with energy E is defined as

$$< |\mathbf{k}_{\|}(E)|^2 >= \frac{\int d\mathbf{k}_i \int d\mathbf{k}_{\|} \, |B_{\mathbf{k}_i}(\mathbf{k}_{\|})|^2 \, |\mathbf{k}_{\|}|^2}{\int d\mathbf{k}_i \int d\mathbf{k}_{\|} \, |B_{\mathbf{k}_i}(\mathbf{k}_{\|})|^2}, \tag{3.15}$$

where $\hbar^2|\mathbf{k}_i|^2/2m = E$. The angle defined by this beam at the screen is given by

$$\Omega(E) = \tan^{-1} \sqrt{\frac{< |\mathbf{k}_{\|}(E)|^2 >}{\frac{2m}{\hbar^2}(E - \phi_{scr}) - < |\mathbf{k}_{\|}(E)|^2 >}}, \tag{3.16}$$

where ϕ_{scr} is the potential of the screen electrode. Note that, Ω is calculated assuming that the motion in the longitudinal and transverse directions are completely decoupled for $z > d$, which may not be valid for the realistic systems. A better way of calculating this angle is to use the quantum mechanical propagation of the electron beam[151] for the emitted wave packet defined by $B_{\mathbf{k}}$. In addition, above defined Ω includes the collimation due to the electric field in the vacuum region $d < z < L$ together with the focusing by the orifice. In the experiments, however, the field collimation may be larger due to large bias between the emitter and screen ($\sim$ 100 V). Therefore, the results presented in this section are not directly comparable with experimental results. We believe, however, that the qualitative results and trends are valid for the realistic cases as well. It should be denoted that the above quantities $< |\mathbf{k}_{\|}|^2 >$ and Ω are functions of the electric field as well. In turn, the collimation angle for the emission can be calculated by integrating Ω over the whole energy range

$$\Theta_c(F) = \int_0^{E_F} dE \; \Omega(E;F) \; J(E;F)/I(F), \tag{3.17}$$

where the energy current density is used as a weighting factor. Finally, the energy spread of the emitted beam can be found as

$$\Delta E(F) = \int_0^{E_F} dE \; (E_F - E) \; J(E;F)/I(F). \tag{3.18}$$

It is in order to use the present formalism to analyze the mechanism for focusing of emitted electron beams from point sources.

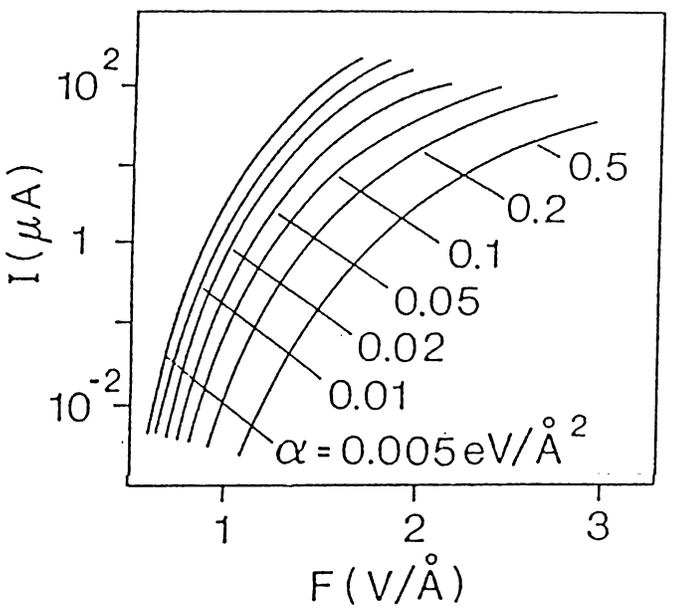

Figure 3.11: Total emission current for emission from a point source: Uniform channel
The emission current as a function of the electric field, the confinement parameter being a parameter.

We first reveal the effect of transverse confinement due to the atomic size of the tip by assuming that α in Equation 3.12 is constant throughout the orifice, i.e., $-r_t < z < d$. From now on we refer this case as a uniform channel. In contrast to the model of García and coworkers,[148,150] the confinement and electric field coexist in the channel. Nevertheless, their effects can be distinguished by varying only α or F at a time. In reality, however, α has to vary as a function of the electric field as found by self-consistent jellium calculations,[149] which is also beyond the scope of our analysis. In Figures 3.11 and 3.12 the total emission current $I(F)$ and collimation angle $\Theta_c(F)$, respectively, calculated for uniform channels are shown. For a given α the current increases and collimation angle decreases with increasing electric field F. This is quite natural since the height and thickness of the tunneling barrier and thus yields relatively larger transmission probability. The collimation gets better as a result of decreasing effective barrier height $\phi_n =$

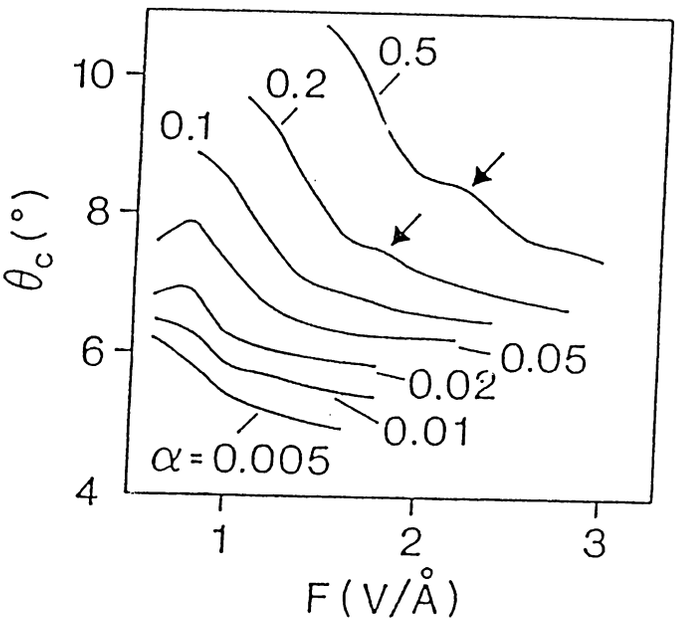

Figure 3.12: Collimation angle for emission from a point source: Uniform channel
The collimation angle as a function of the electric field, the confinement parameter being a parameter

$\max\{\epsilon_n(z) + \phi_m(z)\} - E_F$, thus, increasing the longitudinal momentum $k_n \simeq 2m(E-\phi_n)/\hbar^2$. In fact, the emission angle $\Theta_e(F)$ which is defined just like $\Theta_c(F)$ with $\phi_m(d;F)$ replacing $\phi_{scr}(F)$ has the same qualitative behavior as the latter, only with a larger absolute value. Note that, this decrease in the collimation angle for increasing F is even classically obtainable.

This feature seems to be in contrast with the experimental results[74] which reported increasing spot size (which determines the experimentally measured collimation angle) for increasing potential bias. At this point we have to investigate the experimental measurement method in detail. The screen in the experiments[74] is actually a fluorescent plate and the impact of the electrons yield a bright spot. The brightness of the spot and its size is determined by the properties of the electron beam and the screen as well. For instance, if the flux of electrons impinging on a particular point of the screen is not large enough (larger

than a detection threshold) no fluorescence action will take place and the point will be dark. Therefore, for two electron beams with the same lateral momentum spectra but different total emission currents, the sizes of the corresponding spots on the screen would not be the same. That is, the size of the spot observed on the screen can not be used as a measure of the spread in the momentum space unless the total emission current is kept constant, which is not possible for the real experimental setup. In our case, with increasing electric field the intensity I increases exponentially and Θ_c decreases approximately as a power law. The change in the intensity dominates the experimental detection and for large fields the spot gets both brighter and larger in size.

For increasing α, on the other hand, the tunneling current decreases and collimation angle increases. These observations can be explained easily. The stronger is the lateral confinement, i.e., the larger is α, the higher is the effective barrier, leading to a smaller tunneling probability. The increasing collimation angle is due to the larger lateral momentum content for small α, i.e., $\kappa_n \propto 1/w^2 \propto \alpha^{1/2}$, κ_n being the variance of the lateral momentum for the nth lateral wave function. Since for large α we obtain both large emission current and small collimation angle simultaneously, at this point one may think that the problem of point sources is solved. In fact, for a planar emitter the tunneling current becomes infinite and collimation angle vanishes. However, this view ignores one of the most important feature of the point sources for electron emission, namely the coherence of the emitted electron beams. These electron emitters are thought to be used in areas of application which require phase coherence within the beam across long distances, such as electron holography. Therefore, the size of the emitting source can not be arbitrarily large[154] in order to achieve the coherence at the source. Thus, we have to look for an atomic size emitter with both large enough emission intensity and small enough collimation angle. The small α values included in Figures 3.11 and 3.12, actually, correspond to rather blunt emission sources and are included only for the sake of completeness.

Turning back to Figure 3.12 one observes that the collimation angle Θ_c has an oscillatory behavior for large α values. Surprisingly the energy current density

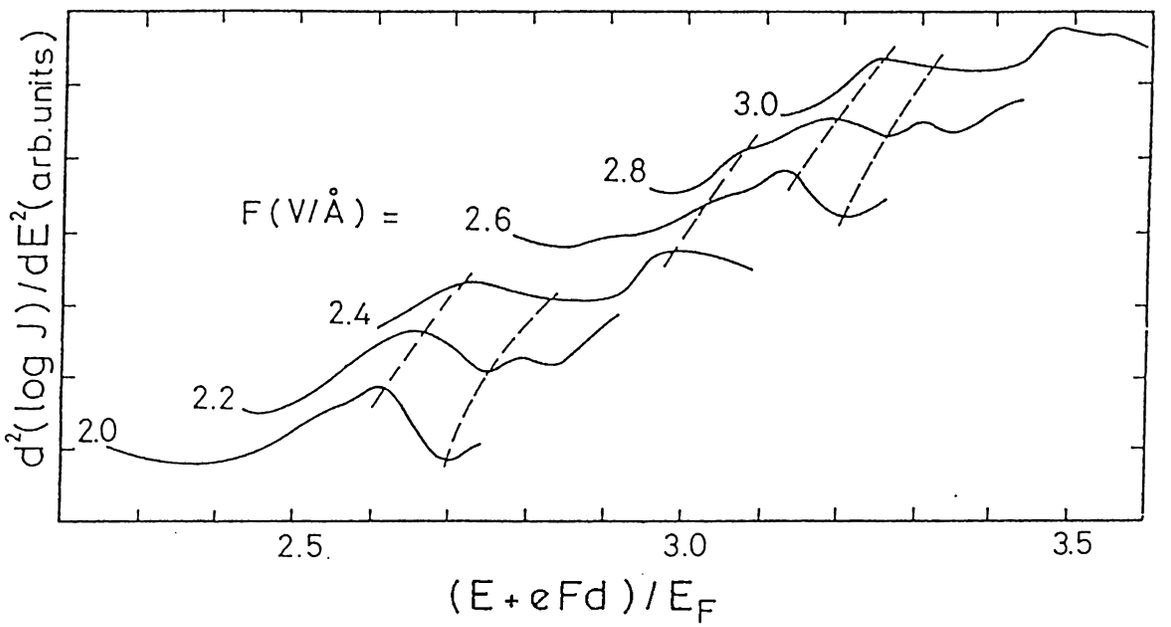

Figure 3.13: Fowler-Nordheim resonances in energy current density
Second derivative of the energy current density as a function of energy as corrected by the electric field, electric field being a parameter.

and total emission current are devoid of any visible structure. Analyzing $J(F, E)$ in detail one finds that it has weak resonances. In Figure 3.13 the second derivative of $\log J(F, E)$ is given for a typical case. The structure of $d^2(\log J)/dE^2$ seems to be shifting by an amount slightly larger than eFd [the change in potential $\phi_m(d; F)$]. This structure may be associated with a Fowler-Nordheim type resonance. For a triangular potential barrier terminated at an abrupt junction such resonances appear as a result of interference of waves between the barrier and the termination junction. The above mentioned difference between the shift of the features and eFd is due to the upward shift of the resonance states with increasing electric field. Although for the strictly 1D case these resonances are strong, because of higher dimensionality and large tunneling barrier these effects are only of minor importance for the present system.

Another important feature of a point source is the subband selection as far as its quasi-1D nature is concerned. In Figure 3.14 the energy current densities for

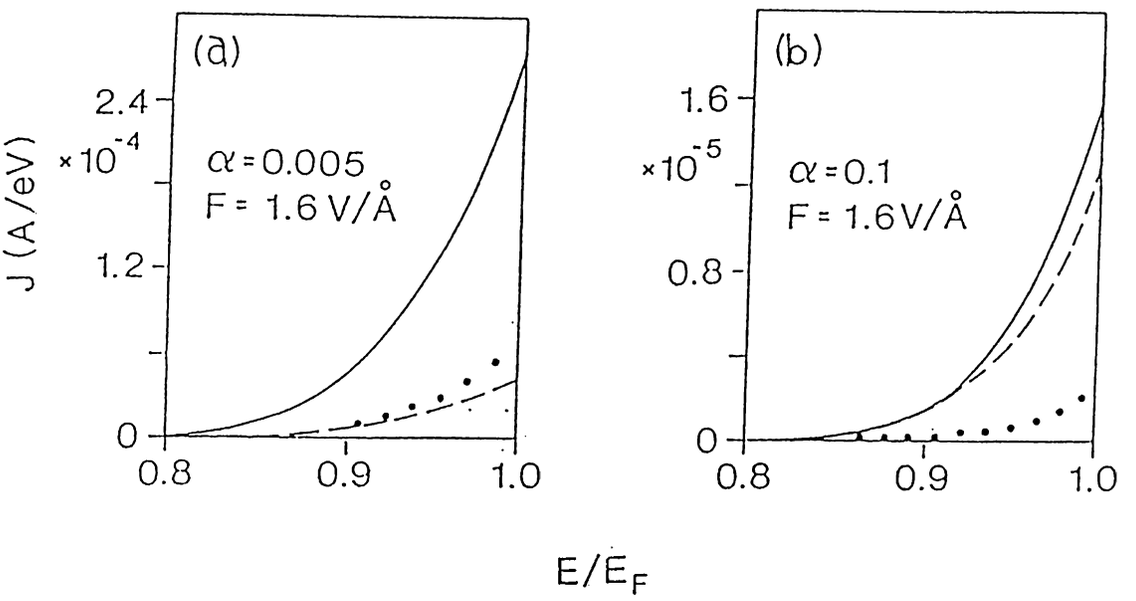

Figure 3.14: Energy current density for uniform channels
The energy current density for uniform channels with (a) $\alpha = 0.005$ eV/Å^2, and (b) 0.1 eV/Å^2. The dashed and dotted lines represent the contribution of the first and second subbands, respectively.

two uniform channels subject to the same electric field are shown. Clearly, for large α the difference between the effective barrier heights for different subbands, $(\phi_n - \phi_{n'})$, is larger as compared to small α. Therefore, the emission is mainly via the lowest subband $n = 0$ for strong lateral confinement. In fact, the collimation angle for strong lateral confinement is close to that determined by using κ_0^2 (expectation value of the lateral momentum squared for the lowest subband) instead of $< |k_{\|}|^2 >$ in Equation 3.16 and for weak confinement it is larger than this value. The contribution of the higher lying subbands for small α may seem to yield an increase in collimation angle Θ_c since the spread of the momentum spectrum increases with subband index. In contrast, decreasing α one still gets an improvement in collimation due to above mentioned relation that $\kappa_n \propto \alpha^{1/2}$. That is, the degrading in the subband selection is dominated by the drop in the emission angle Θ_e. Another observation is that the energy

spread $\Delta E(F)$ increases with increasing electric field and with decreasing α, as expected. However, the dependence on the structural parameters is of only minor importance and ΔE is mainly determined by the electric field F. Consequently, it is not easy to obtain a large emission current and a small energy spread (less than ~ 0.5 eV) simultaneously. This may be one of the problems that will limit the applications of these electron emission sources.

Next we analyze the effects of including a horn-like opening at the emission side, i.e., $z \lesssim d$. As described above this is achieved by changing $w(z)$ linearly between $z = r_h$ and $z = d$. At the end points of the horn the confinement parameters are chosen to be equal to α_t and α_o, respectively. the same kind of tapering may be used at the base of the tip as well. However, the effects of the latter is very small and not interesting. This is due to the fact that at the base of the tip $\phi_m \simeq 0$, and thus first few subbands are well below E_F. Therefore, we concentrate on the horn-like opening to the vacuum. The performance of the horn can be understood by comparing its characteristics with those of uniform channels with $\alpha = \alpha_t$ and α_o. In Figure 3.15 the emission current and in Figure 3.16 the collimation angle for a typical horn geometry are shown. For this system $\alpha_t = 0.25$ eV/Å^2 and $\alpha_o = 0.01$ eV/Å^2, so that the confinement near the tip is strong and it gradually gets weaker as going away from the tip. As seen in Figure 3.15, the total emission current is between those for uniform channels with α_t and α_o. This can be understood by comparing the effective potential profiles for the three configurations. The integrated area of the potential barrier, which approximately determines the transmission probability, for the horn is smaller than that of the uniform channel with α_t and larger than that of uniform channel with α_o. For large F, the emission current approaches to that of uniform channel with α_t, which implies that the transmission is mainly determined by the transmission through the saddle point, where the orifice is the narrowest. This property is reminiscent of the adiabatic evolution[81] which is claimed to exist by Lang and coworkers[149] and is analyzed in detail below. The striking effect of the horn is seen in Figure 3.16, namely a large decrease in Θ_c as compared to the uniform channel with α_t. The collimation angle only 30 Å away from the source

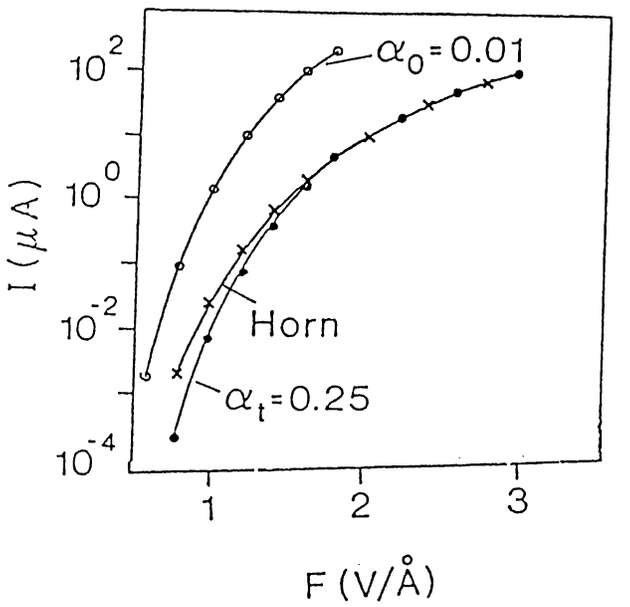

Figure 3.15: Total emission current for emission from a point source: Horn-like opening
The emission current as a function of the electric field for $\alpha_t = 0.25$ eV/Å^2 and $\alpha_o = 0.01$ eV/Å^2. The open and full circles show the emission current for uniform channel with $\alpha = \alpha_t$ and α_o, respectively.

may be as low as $\sim 5°$ for the present configuration for an emission current of ~ 100 μA. The important point is that the size of the source is only a few Å and comparable with the Fermi wavelength, λ_F. On the other hand, the collimation angle is $\sim 7°$ for a uniform channel of the same width. We are not aiming at optimizing the structure, but focus our attention on the underlying physics of this focusing effect.

We analyze the subband selection by the horn geometry as shown in Figure 3.17. Here, the contribution of the lowest lying two subbands ($n = 0$ and 1) to the total emission current are calculated at the exit of the orifice, i.e., $z = d$, are given. Clearly, the filtering of the higher lying subbands is close to, even better than for larger fields, that for the uniform channel with α_o. An analysis of the current distribution along the orifice shows that the contributions of different

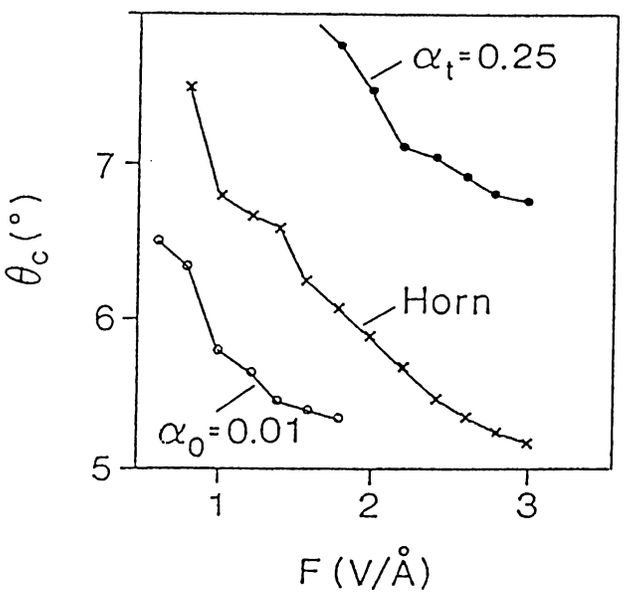

Figure 3.16: Collimation angle for emission from a point source: Horn-like opening
The collimation angle as a function of the electric field for $\alpha_t = 0.25$ eV/Å^2 and $\alpha_o = 0.01$ eV/Å^2. The open and full circles show the collimation angle for uniform channel with $\alpha = \alpha_t$ and α_o, respectively.

subbands does not change appreciable throughout the horn. The absence of scattering among subbands is a property of adiabatic evolution.[81] Thus, one is tempted to think that the horn structure leads to adiabatic transport. However, according to the adiabatic picture[81,149] the subbands evolve through the channel independently so that the total emission current and collimation angle can be written as a sum of individual subband contributions. Explicitly one has

$$I = \sum_n I_n, \tag{3.19}$$

and

$$\Theta_c = \sum_n \Theta_{cn} \simeq \sum_n \vartheta_n I_n / I, \tag{3.20}$$

where ϑ_n is the angle is found by using κ_n^2 instead of $< |k_{||}|^2 >$ in Equation 3.16. These equalities are approximately satisfied for uniform channels for which the

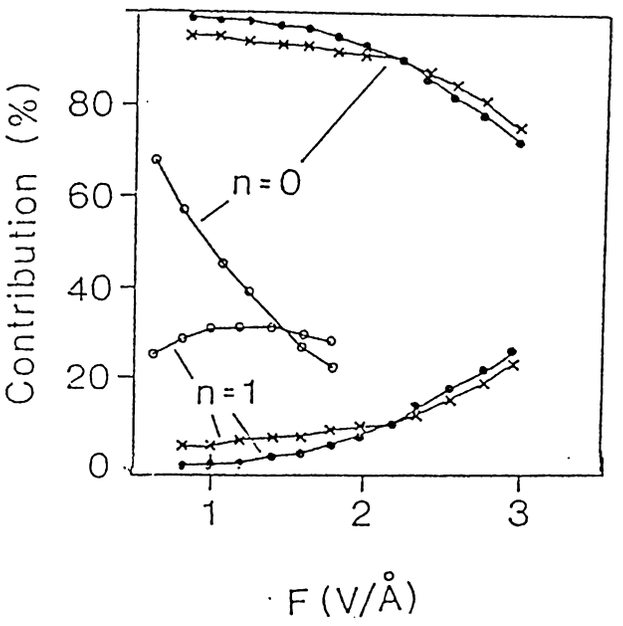

Figure 3.17: Subband selection characteristics for a point source with horn-like opening
The contribution of the first and second subbands to the emission current as a function of the electric field for $\alpha_t = 0.25\ \mathrm{eV}/\mathring{A}^2$ and $\alpha_o = 0.01\ \mathrm{eV}/\mathring{A}^2$. The open and full circles show the collimation angle for uniform channel with $\alpha = \alpha_t$ and α_o, respectively.

subband mixing is small and is possible only at the edges $z = -r_t$ and $z = d$. For large α the difference between the full calculation and diagonal approximation reaches $\sim 0.2^o$ for the collimation angle, since for small w the off-diagonal elements of reflection become appreciable. However, it is clearly not satisfied for the horn structure. This can be seen by comparing Θ_c for the horn geometry and for the uniform channel with α_o together with Figure 3.17. Compared to the uniform channel, the subband selection is improved so that the main contribution to the current is from the lowest subband. Elimination of the higher lying subbands invokes a decrease in Θ_c according to the adiabatic picture, since κ_n increases with n. As depicted in Figure 3.16, in contrast to this expectation, the collimation angle is smaller for the uniform channel with α_o than that is for the horn. For instance, for $F = 1.6\ \mathrm{V}/\mathring{A}$ the full calculation yields $I = 2.04\ \mu\mathrm{A}$

and $\Theta_c = 6.21^o$, whereas the diagonal approximation gives $I = 1.37$ μA and $\Theta_c = 3.59^o$. Thus, the transport through the channel can not be adiabatic as claimed by Lang and coworkers.[149] Although there is no intersubband scattering, the subband mixing causes an interference between different subbands which yields an increase in Θ_c compared to the uniform channel with α_o. That is, the quantum mechanical calculation has to be carried for the whole near-field region, not only for the narrowest part of the channel. Obviously, if adiabatic evolution took place the collimation would be better as compared to the nonadiabatic horn.

To summarize the results of the present calculations one can say that it is possible to to have focused electron beams emerging from atomic size sources with reasonable current values. The essential mechanism for this is the selection of the subbands by the saddle point potential due to the lateral confinement. This way, only the incident waves with small lateral momentum contribute to the emission current. The electric field is responsible for the further collimation of the beam by increasing the longitudinal momentum. A horn-like opening to the vacuum yields an improved collimation, even though the propagation is not adiabatic. Finally we have to add that attention has to be placed on the comparison of experimental results and theoretical findings.

3.4.2 Focusing in a Quantum Point Contact

In the last subsection of this chapter we return to the subject of quantum point contacts studied in Chapter 2. The focusing effect of QPC have important implications in several related problems. The most important one among these is the quenching of Hall effect which was first observed in 1987 by Roukes and coworkers.[155] They found that for ballistic electron waveguides the Hall resistance drops below the classical value for small magnetic field. The quest for understanding the mechanism responsible for this quenching lasted for two years. Recently Baranger and Stone[156] showed that quenching is a result of the collimation by the local widening of the narrow electron waveguides near the cross geometry. Subsequent experiments[157,158] supported this proposition. In addition, the problem of focusing in a QPC draw attention of both theoreticians[102,148]

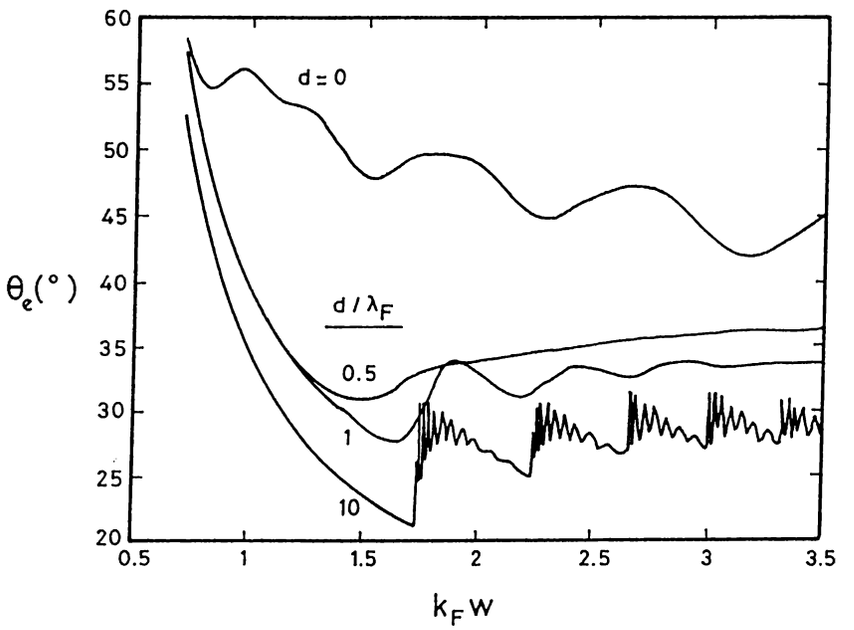

Figure 3.18: Emission angle for quantum point contacts: Uniform constriction The emission angle as a function of width calculated for several values of the constriction length, given in units of λ_F. The curve for $d = 10 \times \lambda_F$ is shifted downwards by 5 ° for clarity. Parabolic confinement.

and experimentalists[159] in past few years. For instance, Beenakker and van Houten[102] explained the nonadditivity of QPC resistances in series and novel magnetotransport effects by use of collimated electron beams. This collimation is observed experimentally[159] as well. In what follows we analyze the reasons and properties of focusing in QPC.

We use a QPC with parabolic confinement as described in Sections 2.2 and 2.3 and we calculate the emission angle Θ_e for the transmitted waves as described in Subsection 3.4.1. First, we use uniform constrictions conductances of which are given in Figure 2.6. In Figure 3.18 the emission angle as a function of the width of the constriction is shown for these uniform QPC. Apparently, Θ_e is quite large for short constrictions, i.e., $d \lesssim \lambda_F/2$. This is due to the large transmission probability for the evanescent modes. That is, incident waves with large lateral momenta can pass through the constriction via tunneling and contribute to the

outgoing waves, increasing the emission angle. For $d \gtrsim \lambda_F$, on the other hand, the tunneling probability is small enough that only incident waves with large enough longitudinal momenta can pass through the constriction leading to focusing of the beam. Note that, collimation does not improve with further increasing the length of the constriction. This is expected, since the emission characteristics is mainly determined by the aperture at the edge of the constriction. An important feature of the θ_e curve for very long constrictions, i.e., $d = 10\ \lambda_F$ in Figure 3.18, is the resonance structure superimposed on a piecewise smooth variation when more than one subband is occupied. The reason for this resonance structure can be sought in Equation 3.15. For example, assume that the lowest lying two subbands are occupied. Then, $B_{\mathbf{k}_i}$ have contributions from both first and second subband. The contribution of the second subband, however, is oscillating due to the interference of the waves in the constriction. Thus, the contribution of the larger lateral momentum part will change between 1/2 (for unity transmission) and a smaller value, yielding to an oscillation of the emission angle. In fact, these resonance positions are in exact match with those of the conductance curve.

Another important property of emission angle is that for large w it approaches to its semiclassical value, which can be calculated by averaging κ_n in a continuum of subbands. It is, however, a little larger than this semiclassical value in this limit (for the present case full calculation gives $\sim 33^\circ$ and the semiclassical value is 30°). This is a result of intersubband interaction and for wide constrictions the correct semiclassical limit can be found by using an integral over the aperture,[102] i.e., by including the wave diffraction effects. In Figure 3.19 the lateral momentum distribution of the emitted wave is shown for two values of w. For the first one only the first subband is occupied and thus the distribution resembles $|\Phi_0(\kappa)|^2$. For the latter, on the other hand, there are five subbands below the Fermi energy. A point to pay attention is that the distribution extends to $\kappa > k_F$, denoting the effect of the states decaying into the 2D EG. Going away from QPC, the contribution of these states to Θ_c decrease. This effect is not taken into account in the preceding section. However, this does not yield drastical changes in $\Theta_c(F)$

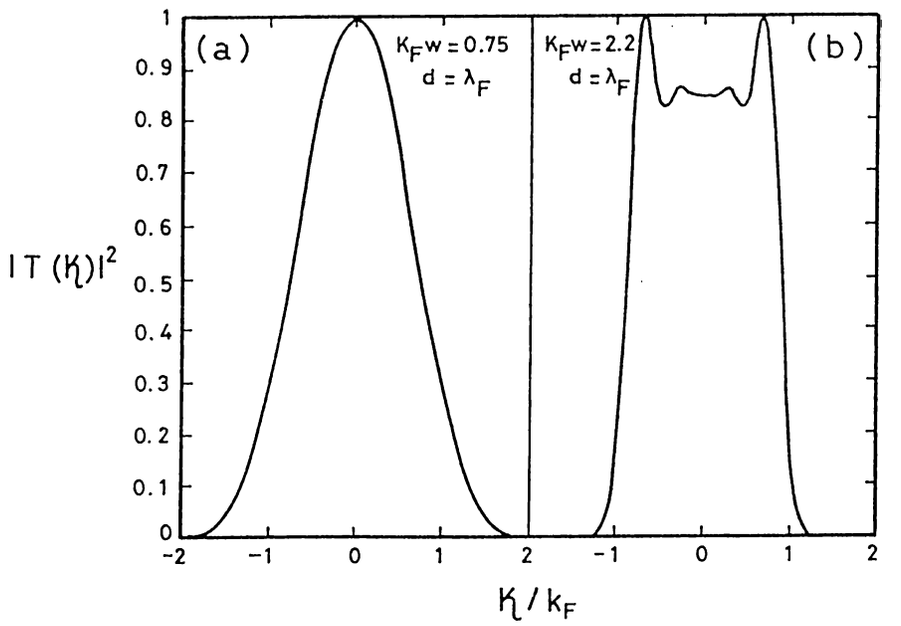

Figure 3.19: Lateral momentum distribution for uniform constriction
The distributions are normalized to have maximum value of unity. There are (a) one, (b) five subbands are below the Fermi level. The length of the constriction is λ_F. Parabolic confinement.

characteristics[‡] due to the subband selection described below. Thus, for the point source the distribution at the aperture to the vacuum is almost zero for $|k_{||}|^2 \lesssim 2m[E_F - \phi_m(d;F)]/\hbar^2$. Note that, the spread of the lateral momentum is approximately the same for the two cases. The reason for this is apparent in Figure 3.18 and was explained by García and coworkers[148] as well. Namely, increasing the width the higher lying subbands begin to be populated and have an increasing effect on the lateral momentum. The increase of the width, however, decreases κ_n. The two effects approximately cancel each other and the emission angle oscillates around its semiclassical value.

Next we consider different constriction configurations in order to investigate

[‡]This statement has the exception of large α and small F values. Still, the effect is not dramatic since at the aperture to the vacuum the first subband is still below E_F and the decaying states become propagating for $z \gtrsim d$ due to the electric field.

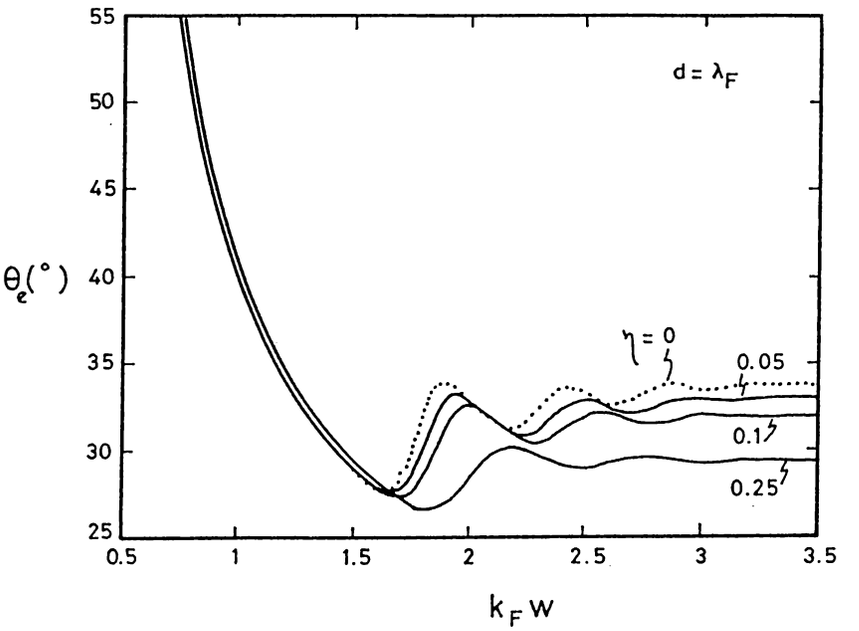

Figure 3.20: Focusing in a quantum point contact: Barrier in the constriction The emission angle as a function of width calculated for several forms of saddle point potential as given by Equation 2.79. The length of the constriction is λ_F. The dotted curve is for the uniform constriction.

the agents responsible for focusing. The existing studies[102,150] point out two main mechanisms of focusing, namely, the subband filtering and horn-effect. The essence of subband filtering is to eliminate the higher lying subbands by including a potential barrier in or at the exit of the constriction for large w. Thus, only the low index subbands will contribute the current and large lateral momentum components are reflected back. Adiabatic evolution of states, on the other hand, performs the same scenario by means of the effective potential. That is, the narrowest part of the constriction allows only the low index subbands to propagate and expanding the width smoothly the lateral momentum is decreased towards the exit of the constriction. It has to be reminded that in the previous subsection both of these mechanisms played a role in the focused field emission.

In Figures 3.20 and 3.21 the emission angle characteristics for constrictions having barriers and narrowings, respectively, are shown. As expected, both

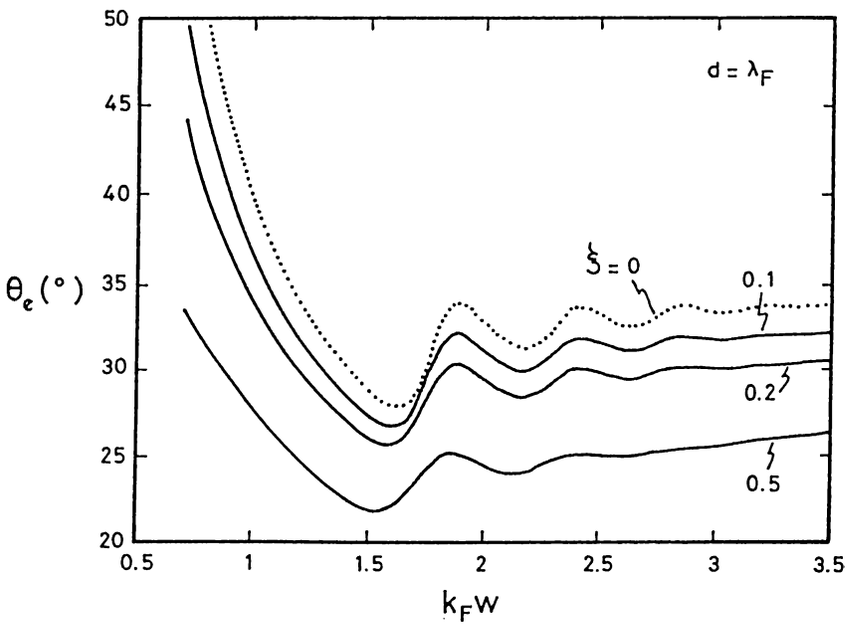

Figure 3.21: Focusing in a quantum point contact: Tapering of the constriction The emission angle as a function of width calculated for several forms of tapering as given by Equation 2.80. The length of the constriction is λ_F. The dotted curve is for the uniform constriction. Parabolic confinement.

configurations yield collimation so that the emission angle decreases as compared to the uniform constriction. For the barrier in the constriction case the semiclassical value of Θ_e is smaller than that for the uniform constriction. This is a result of reduced electron energy in the constriction. That is, in order to have propagation in a subband the energy of the incoming waves has to be greater than the subband energy plus the potential barrier height. Therefore, the oscillatory features of the emission angle curve are shifted to larger w values and although κ_n is the same with the uniform channel Θ_e is decreased. This result is in complete agreement with the semiclassical explanation.[102] For the tapered constriction, however, the semiclassical value is the same as the uniform channel. This is due to the nonadiabatic evolution along the constriction. That is, for large w values the higher lying subbands start to get occupied as a result of tunneling and lateral momentum increases. This effect disappears when the barrier and

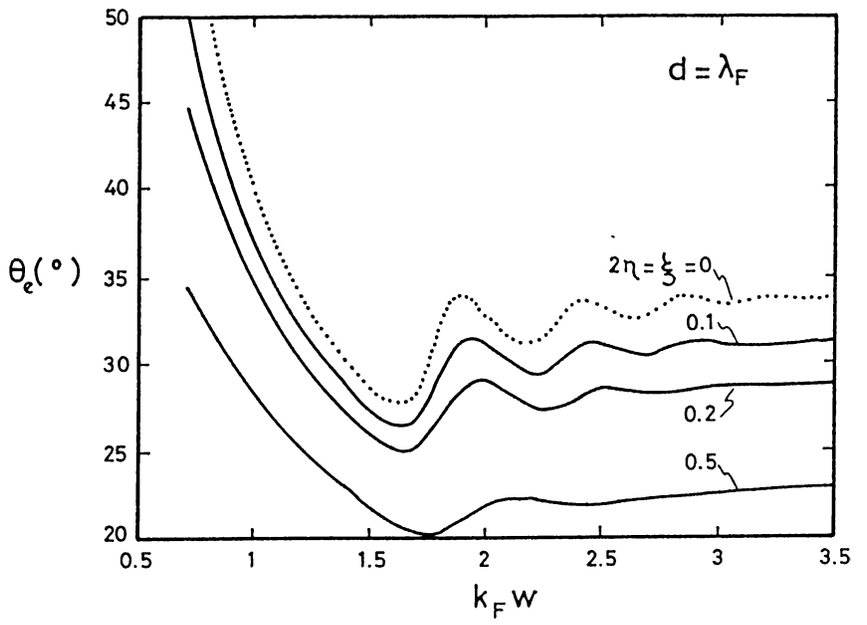

Figure 3.22: Focusing in a quantum point contact: Barrier and tapering coexist The emission angle as a function of width calculated for several forms of tapering and saddle point potential together. The length of the constriction is λ_F. The dotted curve is for the uniform constriction. Parabolic confinement.

tapering coexist in the constriction as shown in Figure 3.22.

Finally we want to emphasize an important difference between the quantum point contacts investigated in this subsection and the point sources analyzed in the preceding subsection. As seen in Figures 3.20, 3.21 and 3.22, for the quantum limit of the QPC, i.e., when only a few subbands are occupied, the semiclassical argumentation does not apply *per se*. This is originating from several effects. To give an example, one may observe that for a nearly pinched off QPC inserting a barrier in the constriction makes the focusing worst. Therefore, one has to carry out full-quantum mechanical calculations to understand the properties of such point sources. A similar comment was put forward by Roukes and coworkers[160] about the semiclassical explanation of the quenched Hall effect[161] for the cross bar systems with only a few occupied subbands.

3.A Self-consistent Field Pseudopotential Calculations for Al tip-Al Sample System

Ciraci and coworkers[134–136] carried out extensive calculations on STM of graphite and Al surfaces in order to resolve the tip-sample interaction. These are SCF pseudopotential calculations of great accuracy and *ab-initio* nature. Still the method has two serious limitations. These are due to the repeating tip structure and applicability of the local density approximation (LDA).

Since the pseudo-wave functions, which are the eigenstates of the pseudo-hamiltonian, are expressed as

$$\varphi_n(\mathbf{k},\mathbf{r}) = \sum_{\mathbf{G}} a_n(\mathbf{k}+\mathbf{G})\; e^{i(\mathbf{k}+\mathbf{G})\cdot\mathbf{r}}, \tag{3.21}$$

in terms of a plane wave basis set, most of the calculations are carried out in momentum space. This brings the necessity of repating the tip structure, forming an array of tips. This can be tolerated if the tip-tip distance is large enough. However, increasing the tip-tip distance, that is increasing the size of the unit cell, the size of the Brillouin zone decreases. Therefore, in order to have a fixed energy cut-off, the number of plane waves that is used in the calculations has to be increased. This number determines both the computation time and memory size. In this respect the tip-tip distance can not be increased arbitrarily. In these calculations LDA is used to obtain the self-consistent field. It is known that at low electron density this approximation is not valid. Moreover, the tails of the wave functions far from the surface can not be represented well. This becomes important for wide barriers. Due to these constraints the tip-sample distance can not be increased too much.

In the most recent calculations carried out by Ciraci and coworkers[136] a (3×3) array of tips connected to the (111) surface of Al is used, yielding to ~ 9 $\AA$ tip-tip distance. The range of the tip-sample distance, on the other hand, had to lie in the interval $3-4 \lesssim h \lesssim 14$ a.u. Here h denotes the interatomic distance and not the distance between the jellium edges, i.e., d. The kinetic energy cut-off is chosen to be $|\mathbf{k}+\mathbf{G}|^2 < 5$ Ry. leading to $\sim$ 2000 plane waves for the

basis set. The self-consistency criterion is 10^{-6} Ry. deviation of the potential energy between two consecutive iterations. The Schrödinger equation for the pseudo-hamiltonian is solved at 15 special points in the irreducible part of the Brillouin zone. The Ceperly-Alder exchange-correlation potential with the non-local, norm-conserving pseudopotential given by Hamann and coworkers is used in the calculation. Calculation of the total energy and forces is not feasible since the size of the unit cell is changing according to the tip-sample distance. The related references and details can be found in References 134 and 136.

The arrangement of the atoms is different for the ontop- and hollow-site position of the tip for these calculations as compared to those given in Reference 135. This is due to a more realistic tip model. The apex of the tip is chosen to consist of four atoms in the form of a pyramid. For the ontop-site, this pyramid has the $C'A'$ arrangement in terms of the (111) layers of Al. It is put on top of an $ABCAB$ arrangement of Al layers as a continuation. This five layers act both as the base of the tip and sample. Since the atomic arrangement is repated in all three dimensions the A' atom has the A layer of the next unit cell on top of it. By changing the size of the unit cell in this direction it is possible to adjust the tip-sample distance. The thickness of the slab has to be large enough in order to prevent the mixing between the states characterizing the tip and sample. This is the reason for using five layers, thus having 49 atoms in the unit cell. For the hollow-site, on the other hand the lowest A layer is removed so that the A' atom has the B layer on top of it. This unit cell contains 40 atoms.

In Figure 3.23 the charge distribution and self-consistent potential is shown for the ontop-site position of the tip. The results of the SCF calculations showed that the jellium approximation for the longitudinal potential and a parabolic confining potential approximately represents the self-consistent potential.

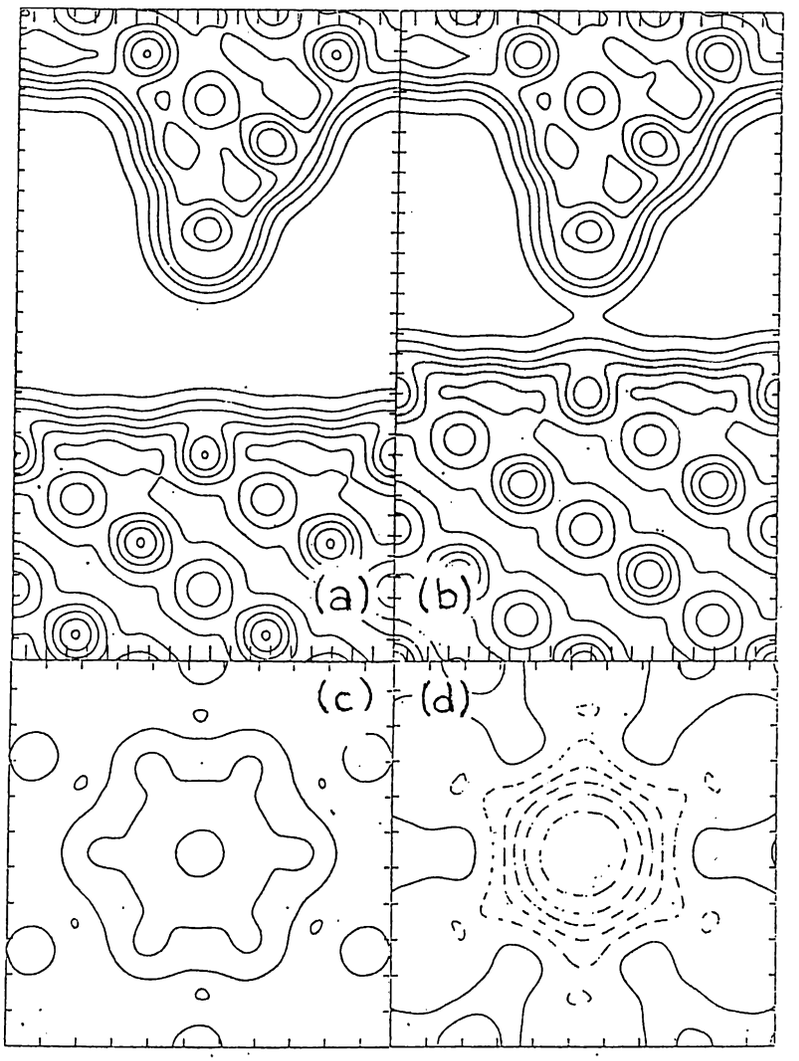

Figure 3.23: Self-consistent charge density and potential obtained for Al tip-Al sample system

The SCF charge density for the ontop-site position of the tip with the tip-sample internuclear spacing of (a) 11 and (b) 8 a.u. The contour spacing is 0.005 electrons/a.u.$^{-3}$. The SCF potential in the bisecting plane for (c) 11 and (d) 8 a.u. are also shown. The contour spacing is 0.02 Ry. and the minimum value of the potentials are -0.10 and 0.11 Ry., respectively, measured from the Fermi level. The dashed contours denote potential below the Fermi level. The ticks on the axes are 1 a.u. apart.

Chapter 4

Nonequilibrium Theory of Quantum Transport

4.1 Equations of Motion in Steady-State

As discussed in Section 1.4, the transport measurements are nonequilibrium processes. Although in Chapters 2 and 3 techniques for systems in equilibrium are used to understand transport and tunneling in mesoscopic structures, it is inevitably necessary to devise a nonequilibrium theory in order to get a deeper understanding.

Nonequilibrium state has been one of the most interesting problems in both classical and quantum physics. Theoretical methods for dealing with nonequilibrium processes have been formulated in early 1960's.[162,163] In fact, the formalisms of Kadanoff and Baym[162] and Keldysh[163] are equivalent to each other, despite their formal differences. The underlying motivation of these methods is to use different Green's function for electrons and holes (in the same band). For equilibrium states, the properties of one type of charge carriers can be determined by making use of the other. However, for nonequilibrium dynamics and kinetics it is necessary to take into account both of them independently. Consequently, the number of independent Green's functions is three for the nonequilibrium case, compared to one for the equilibrium. Two of these incorporate information

about the dynamics of the electrons and holes, and the third one is related to the statistical properties of the system.[163] Thus, it is possible to derive a quantum transport equation[162,163] for this last Green's function which at the classical limit reduce to Boltzmann equation. Since the subject matter of this chapter is the nonequilibrium transport in mesoscopics, we rather use the equations of motion as they stand. The formalism developed by Keldysh[163–165] has the appealing property that it is possible to use the well-known equilibrium Green's function formalism in a matrix form. This simplifies the notation a lot, thus we adopt the Green's function technique of Keldysh[‡].

The Dyson equation in the Keldysh formalism reads[163,165,166]

$$[i\frac{\partial}{\partial t} - \hat{H}_o(\mathbf{r})]\tilde{\mathbf{G}}(\mathbf{r},t;\mathbf{r}',t') = \delta(\mathbf{r}-\mathbf{r}')\delta(t-t')\tilde{\mathbf{I}} + \int d\mathbf{r}''dt''\ \tilde{\mathbf{\Sigma}}(\mathbf{r},t;\mathbf{r}'',t'')\tilde{\mathbf{G}}(\mathbf{r}'',t'';\mathbf{r}',t'), \quad (4.1)$$

where $\hat{H}_o$ is the Hamiltonian of the system in the absence of the self-energy corrections due to both elastic and inelastic scattering included in $\tilde{\mathbf{\Sigma}}$. $\tilde{\mathbf{I}}$ is the (2×2) identity matrix. $\tilde{\mathbf{G}}$ is the matrix of Green's functions, and it includes the electron, hole, time ordered and anti-time ordered propagators in the form

$$\tilde{\mathbf{G}} = \begin{bmatrix} G^T & -G^< \\ G^> & -G^{\bar{T}} \end{bmatrix}. \quad (4.2)$$

These Green's function are defined as

$$G^<(\mathbf{r},t;\mathbf{r}',t') = i<\hat{\psi}^\dagger(\mathbf{r}',t')\hat{\psi}(\mathbf{r},t)>, \quad (4.3)$$

$$G^>(\mathbf{r},t;\mathbf{r}',t') = -i<\hat{\psi}(\mathbf{r},t)\hat{\psi}^\dagger(\mathbf{r}',t')>, \quad (4.4)$$

$$G^T(\mathbf{r},t;\mathbf{r}',t') = \theta(t-t')G^>(\mathbf{r},t;\mathbf{r}',t) + \theta(t'-t)G^<(\mathbf{r},t,\mathbf{r}',t), \quad (4.5)$$

$$G^{\bar{T}}(\mathbf{r},t;\mathbf{r}',t') = \theta(t'-t)G^>(\mathbf{r},t;\mathbf{r}',t') + \theta(t-t')G^<(\mathbf{r},t;\mathbf{r}',t'), \quad (4.6)$$

where $\hat{\psi}$ and $\hat{\psi}^\dagger$ are the annihilation and creation field operators, respectively. The angular brackets denote averaging over all available states of the system. It

[‡]A brief summary of Keldysh technique is given in Appendix 4.A. We use $\hbar = 1$ throughout this chapter.

is convenient to make use of the retarded and advanced Green's functions as well

$$G^R(\mathbf{r},t;\mathbf{r}',t') = \theta(t-t')\,[G^>(\mathbf{r},t;\mathbf{r}',t') - G^<(\mathbf{r},t;\mathbf{r}',t')], \quad (4.7)$$

$$G^A(\mathbf{r},t;\mathbf{r}',t') = \theta(t'-t)\,[G^>(\mathbf{r},t;\mathbf{r}',t') - G^<(\mathbf{r},t;\mathbf{r}',t')]. \quad (4.8)$$

Note that, these Green's functions have similar but nonidentical meanings as those used for systems at equilibrium.

It is possible to derive a quantum Boltzmann equation[162,163,166] starting from Equation 4.1. This can be done by transforming it to the center-of-mass and relative coordinates and then taking the Fourier transform with respect to the relative ones. This way a partial differential equation in time and space coordinates is found for the Wigner distribution function[162] $f(\mathbf{r},t;\mathbf{k},E)$. As compared to the classical Boltzmann equation, energy E appears as a variable of the distribution function in addition to $\mathbf{r}$, $\mathbf{k}$ and t. This is due to the fact that in quantum scale the relation connecting these four variables is correct only in the form of expectation values, that is $E \neq |\mathbf{k}|^2/2m + V(\mathbf{r},t)$. The quantum Boltzmann equation or other quantum kinetic equations were shown[166] to be quite appropriate to study the properties of systems out of equilibrium if the equilibrium system is uniform or the momentum variable $\mathbf{k}$ has a well defined meaning. However, for the mesoscopic devices we are considering, $\mathbf{k}$ does not have a well defined meaning in this context. Still, it is possible to take the Fourier transform over the relative time coordinate to obtain

$$\tilde{\mathbf{G}}(\mathbf{r},\mathbf{r}';E,T) = \int d\tau e^{iE\tau}\tilde{\mathbf{G}}(\mathbf{r},T+\tau/2;\mathbf{r}',T-\tau/2), \quad (4.9)$$

with $T = (t+t')/2$ and $\tau = (t-t')$. Since we are concerned with the steady-state behavior of the system and ignore the transients, the time variable T becomes irrelevant. Thus we are left with the nonequilibrium Green's function $\tilde{\mathbf{G}}(\mathbf{r},\mathbf{r}';E)$. Transforming Equation 4.1 accordingly one finds that

$$[E - \hat{H}_o(\mathbf{r})]\tilde{\mathbf{G}}(\mathbf{r},\mathbf{r}';E) = \delta(\mathbf{r}-\mathbf{r}')\tilde{\mathbf{I}} + \int d\mathbf{r}''\ \tilde{\mathbf{\Sigma}}(\mathbf{r},\mathbf{r}'';E)\tilde{\mathbf{G}}(\mathbf{r}'',\mathbf{r}';E). \quad (4.10)$$

This equation was used by Datta[167] to derive a quantum kinetic equation. He claimed that only the diagonal elements of the Green's function is involved in

the final transport equation. This statement, however, was incorrect due to the appearance of the off-diagonal elements in the kernel of the transport equation. This is very natural, since the spatial correlations represented by the off-diagonal elements are an integral part of the quantum mechanical nature of the system. This become important, for example, when the nonlocality of quantum transport is analyzed.

We rather convert the differential equation in Equation 4.10 to an integral equation first. This can be done by considering the equation satisfied by the noninteracting Green's function $\tilde{\mathbf{G}}_o$

$$[E - \hat{H}_o(\mathbf{r})]\tilde{\mathbf{G}}_o(\mathbf{r}, \mathbf{r}'; E) = \delta(\mathbf{r} - \mathbf{r}')\tilde{\mathbf{I}}. \tag{4.11}$$

$\tilde{\mathbf{G}}_o$ is used as a kernel to solve Equation 4.10. This yields a perturbative expansion for $\tilde{\mathbf{G}}$

$$\tilde{\mathbf{G}}(\mathbf{r}, \mathbf{r}'; E) = \tilde{\mathbf{G}}_o(\mathbf{r}, \mathbf{r}'; E) + \int d\mathbf{r}'' d\mathbf{r}''' \; \tilde{\mathbf{G}}_o(\mathbf{r}, \mathbf{r}''; E)\tilde{\Sigma}(\mathbf{r}'', \mathbf{r}'''; E)\tilde{\mathbf{G}}(\mathbf{r}''', \mathbf{r}'; E), \tag{4.12}$$

which is the Dyson equation in the integral form.

We wish to discuss the differences between the differential and integral forms of the Dyson equation in the context of quantum transport. In measurement experiments the measured values of voltage and current appear as some kind of boundary condition of the system. For example, the current going into a particular probe is a boundary condition for the incoming and outgoing waves in that probe. The voltage measurements are a bit more complicated[76] since it is not possible to define an electrochemical potential due to the nonequilibrium nature. Still, the potentials of the voltage probes appear as boundary conditions for the electron density deep in the probes. For the differential form of the Dyson equation these boundary conditions have to be satisfied by the Green's function $\tilde{\mathbf{G}}$ together with the differential equation, Equation 4.10. In contrast, the integral form of the Dyson equation may be solved together with the boundary conditions by including these boundary conditions within the noninteracting Green's function $\tilde{\mathbf{G}}_o$. This simplifies the problem to a great extent and presents

a more realistic approach to quantum transport. In Sections 4.3 and 4.4 this method is exemplified in detail.

Returning to Equation 4.12 one observes that the resulting scalar equations are[166]

$$G^R = G_o^R\,(1+\Sigma^R G^R), \tag{4.13}$$

$$G^A = G_o^A\,(1+\Sigma^A G^A), \tag{4.14}$$

$$G^< = (1+G^R\Sigma^R)\,G_o^<\,(1+\Sigma^A G^A) + G^R\,\Sigma^<\,G^A, \tag{4.15}$$

where we suppress the coordinates and multiplication stands for an integration over the space, 1 denoting the δ-function. These are the equations that we use to calculate the full Green's functions and related quantities. However, as they stand they are very hard to deal with. Therefore, we use some approximations to simplify the system under consideration without destroying the physical essence.

First, we consider a strictly 1D system. This corresponds to an electron waveguide in the strong quantum limit, i.e., with a single occupied subband. In fact, as we show in Section 2.4 there are certain similarities between scattering in quasi-1D and strictly 1D systems. Nevertheless, the novel features (such as enhanced backscattering) are not observable for the 1D system. Thus, we use a 1D model in expense of loosing such properties unique to quasi-1D systems. In fact, the generalization of the formalism to quasi-1D (i.e., multisubband case) is straight forward as described below. Secondly, we assume that the whole interaction through the system is represented by a number of discrete, local, and uncorrelated scattering centers. That is, the self-energy is written as

$$\tilde{\Sigma}(x,x';E) = \sum_i \tilde{\Sigma}_i(E)\delta(x-x')\delta(x-x_i), \tag{4.16}$$

where i is the index of the scattering center, $\tilde{\Sigma}_i$ is the corresponding self-energy and x_i is its position. $\tilde{\Sigma}_i(E)$ is determined by the scattering process that takes place on the ith center. This self-energy expression is appropriate, for instance, for phonons localized around defects such as impurities, interfaces, etc. The multisubband generalization of the self-energy will be in the form of supermatrices $[\tilde{\Sigma}_i]_{nm}$, where i denotes the scatterer and n and m are the subband

indices. All the following formulation has the same for in terms of these super-matrices for the quasi-1D case.

Using the self-energy expression given in Equation 4.16, Equations 4.15 and 4.13 take the form

$$G^R(x,x';E) = G_o^R(x,x';E) + \sum_i G_o^R(x,x_i;E)\Sigma_i^R(E)G^R(x_i,x';E), \quad (4.17)$$

$$\begin{aligned} G^<(x,x';E) &= G_o^<(x,x';E) + \sum_i G_o^R(x,x_i;E)\Sigma_i^R(E)G^<(x_i,x';E) \\ &+ \sum_i G_<(x,x_i;E)\Sigma_i^A(E)G_o^A(x_i,x';E) \\ &+ \sum_i G^R(x,x';E)\Sigma_i^<(E)G^A(x_i,x';E) \\ &+ \sum_{ij} G^R(x,x';E)\Sigma_i^R(E)G_o^<(x_i,x_j;E)\Sigma_j^A(E)G_o^A(x_i,x';E), \end{aligned} \quad (4.18)$$

respectively. In the remaining part of this section we solve these two equations in order to find $G^<(x,x';E)$ and calculate the electron density and current therefrom.

Taking x and x' at the position of scatterers, Equation 4.17 yields

$$G^R(x_i,x_j;E) = G_o^R(x_i,x_j;E) + \sum_k G_o^R(x_i,x_k;E)\Sigma_k^R(E)G^R(x_k,x_j;E). \quad (4.19)$$

which can be cast in a matrix form with the solution

$$\tilde{\mathbf{G}}^R(E) = [\tilde{\mathbf{I}} - \tilde{\mathbf{G}}_o^R(E)\tilde{\mathbf{\Sigma}}^R(E)]^{-1}\tilde{\mathbf{G}}_o^R(E), \quad (4.20)$$

where the elements of these matrices read $G_{ij}^R(E) = G^R(x_i,x_j;E)$, $I_{ij} = \delta_{ij}$ and $\Sigma_{ij}^R(E) = \delta_{ij}\Sigma_i^R(E)$, respectively. These matrices has to be differentiated from those used in the full Dyson equation, Equations 4.10 and 4.12. The indices there correspond to the type of the Green's function and here to the scattering centers. Similarly, one can show that

$$G^R(x_i,x';E) = \sum_j [\tilde{\mathbf{I}} - \tilde{\mathbf{G}}_o^R(E)\tilde{\mathbf{\Sigma}}(E)]_{ij}^{-1} G_o^R(x_j,x';E), \quad (4.21)$$

which yields the solution for the retarded Green's function

$$G^R(x,x';E) = G_o^R(x,x';E) + \sum_{ij} G_o^R(x,x_i;E)[\tilde{\mathbf{I}} - \tilde{\mathbf{G}}_o^R(E)\tilde{\mathbf{\Sigma}}(E)]^{-1}_{ij} G_o^R(x_j,x';E). \tag{4.22}$$

Equation 4.18 is solved following similar steps. The result can be written as

$$\begin{aligned} G^<(x,x';E) = G_o^<(x,x';E) &+ \sum_{ij} G_o^R(x,x_i;E)\{\Sigma_*^R\}_{ij}(E)G_o^<(x_j,x';E) \\ &+ \sum_{ij} G_o^<(x,x_i;E)\{\Sigma_*^A\}_{ij}(E)G_o^A(x_j,x';E) \\ &+ \sum_{ij} G_o^R(x,x_i;E)\{\Sigma_*^<\}_{ij}(E)G_o^A(x_j,x';E), \end{aligned} \tag{4.23}$$

where the reducible self-energy $\tilde{\Sigma}_*$ is defined as

$$\tilde{\Sigma}_*^R(E) = [\tilde{\mathbf{I}} - \tilde{\Sigma}^R(E)\tilde{\mathbf{G}}_o^R(E)]^{-1}\tilde{\Sigma}^R(E), \tag{4.24}$$

$$\tilde{\Sigma}_*^A(E) = \tilde{\Sigma}^A(E)[\tilde{\mathbf{I}} - \tilde{\mathbf{G}}_o^A(E)\tilde{\Sigma}^A(E)]^{-1}, \tag{4.25}$$

$$\begin{aligned} \tilde{\Sigma}_*^<(E) = {} & [\tilde{\mathbf{I}} - \tilde{\Sigma}^R(E)\tilde{\mathbf{G}}_o^R(E)]^{-1}\{\tilde{\Sigma}^<(E) + \tilde{\Sigma}^R(E)\tilde{\mathbf{G}}_o^<(E)\tilde{\Sigma}^A(E)\} \\ & \times[\tilde{\mathbf{I}} - \tilde{\mathbf{G}}_o^A(E)\tilde{\Sigma}^A(E)]^{-1}. \end{aligned} \tag{4.26}$$

We formulated a method to find the nonequilibrium Green's functions for a given distribution and strength of scatterers. Now we have to express measurable quantities in terms of these Green's functions. In fact, $G^<$ and $G^>$ include all the information concerning the electrons and holes, respectively. For example the electron density is

$$n(x;E) = < \hat{\psi}^\dagger(x;E)\hat{\psi}(x;E) >, \tag{4.27}$$

which is nothing but $-iG^<(x,x:E)$. Note that, $n(x;E)$ is not the usual electron density but the local density of the occupied electronic states. We refer it using the former term in the following, for the sake of briefity. From Equation 4.23 one finds

$$\begin{aligned} n(x;E) = n_o(x;E) &+ \sum_{ij} 2 \text{ Im } [G_o^R(x,x_i;E)\{\Sigma_*^R\}_{ij}G_o^<(x_j,x:E)] \\ &- \sum_{ij} i\,[G_o^R(x,x_i;E)\{\Sigma_*^<\}_{ij}G_o^A(x_j,x:E)], \end{aligned} \tag{4.28}$$

where we made use of $(\tilde{\mathbf{G}}_o^<)^\dagger = -\tilde{\mathbf{G}}_o^<$, $(\tilde{\mathbf{G}}_o^R)^\dagger = \tilde{\mathbf{G}}_o^A$, and $(\tilde{\mathbf{\Sigma}}^<)^\dagger = -\tilde{\mathbf{\Sigma}}^<$, $(\tilde{\mathbf{\Sigma}}^R)^\dagger = \tilde{\mathbf{\Sigma}}^A$. The first two of these equalities follow from the definition of the Green's functions. The last two, on the other hand, can be derived from the Dyson equation requiring that the same relations apply for the interacting Green's functions as well. Since the noninteracting and full Green's functions can be written in terms of the corresponding field operators,[163,164] this requirement is quite natural. In Equation 4.28, $n_o(x;E)$ denotes the electron density in the absence of interaction.

The current (the local energy current density, to be precise), on the other hand is given by

$$j(x;E) = \frac{1}{2im} < \left[\frac{\partial}{\partial x}\hat{\psi}^\dagger(x;E)\hat{\psi}(x';E) - \hat{\psi}^\dagger(x;E)\frac{\partial}{\partial x'}\hat{\psi}(x';E)\right]_{x'=x} >, \quad (4.29)$$

which in turn reduces to

$$j(x;E) = \frac{-1}{2m}\left[\frac{\partial}{\partial x} - \frac{\partial}{\partial x'}\right]_{x'=x} G^<(x,x';E). \quad (4.30)$$

Using Equation 4.23 and the relations mentioned above one finds

$$\begin{aligned} j(x;E) = j_o(x;E) \;+\; & \sum_{ij}\frac{-1}{m}\,\mathrm{Re}\left[\left(\frac{\partial}{\partial x}G_o^R(x,x_i;E)\right)\{\Sigma_*^R\}_{ij}G_o^<(x_j,x:E)\right] \\ - \; & \sum_{ij}\frac{-1}{m}\,\mathrm{Re}\left[G_o^R(x,x_i;E)\left\{\{\Sigma_*^R\}_{ij}\left(\frac{\partial}{\partial x}G_o^<(x_j,x:E)\right)\right.\right. \\ & \left.\left.\{\Sigma_*^<\}_{ij}\left(\frac{\partial}{\partial x}G_o^A(x_j,x:E)\right)\right\}\right], \end{aligned} \quad (4.31)$$

where $j_o(x;E)$ is the current for the noninteracting system.

A question about the current is if the continuity equations are satisfied. Clearly, for the noninteracting system the current is constant throughout the device

$$\frac{\partial}{\partial x}j_o(x;E) = 0, \quad (4.32)$$

due to the steady-state nature. However, the self-energy $\tilde{\mathbf{\Sigma}}$ may contain terms which violate the current conservation. For example, a current probe, by its very definition, draws current from the system (or injects current into the system).

Therefore, the right-hand side of Equation 4.32 may not be not zero for $j(x;E)$. The conservation equation satisfied by $j(x;E)$ is derived from Equation 4.31 and after some algebra it reads

$$\frac{\partial}{\partial x}j(x;E) = \sum_i 2i\delta(x-x_i)[\mathrm{Im}\{G^R(x,x_i;E)\}\Sigma_i^<(E) - G^<(x,x_i;E)\mathrm{Im}\{\Sigma_i^R(E)\}]. \tag{4.33}$$

That is, conservation of current may be violated only at the scattering centers and this happens only if the number of electrons scattered is not equal to the number of holes scattered. That is, only if part of the current goes into or comes out of the scatterer, as expected. For a point scatterer this is not understandable, since it can not absorb charge carriers in the steady-state. However, an inelastic scatterer may represent a current probe, which is connected to a reservoir (which is not within the system described by the Hamiltonian $\hat{H}$) that supplies the necessary number of charge carriers to violate the conservation in steady-state. In Section 4.3 we use this property for building model current probes. An important quantity is the discontinuity in current $\Delta j_i(E) = j(x_i^+;E) - j(x_i^-;E)$, which is found as

$$\Delta j_i(E) = [n(x_i;E)\{i\Sigma_i^>(E)\} + p(x_i;E)\{i\Sigma_i^<(E)\}], \tag{4.34}$$

where $p(x;E)$ is the local density of the empty electronic states (we refer it as the hole density). Note that, given an arbitrary set of scatterers with $\tilde{\Sigma}_i(E)$, Equation 4.34 yields $\Delta j_i \neq 0$. However, if there are no current probes this contradicts the physical reality. Therefore, $\tilde{\Sigma}_i(E)$ has to be found self-consistently in order to satisfy the conservation of current. This self-consistency may be achieved by either varying the strength of inelastic scattering, or changing the self-consistent potential or doing both, depending on the system. If the electron and hole density is large enough, then the screening will be effective and self-consistent potential variations will be small. In the following section this kind of self-consistency is used, that is, the strength of inelastic scattering is varied according to the density of carriers.

In Equation 4.34, the first and second terms in the brackets correspond to the absorption of electrons and holes, respectively, of energy E absorbed by the ith

scatterer. Thus, $\{i\Sigma^{>}(E)\}$ is the rate of electron scattering into the scattering center out of energy E assuming it is initially full and $\{-i\Sigma^{<}(E)\}$ is the rate of hole scattering into the scattering center out of energy E (or equivalently electron scattering out of scattering center into energy E) assuming it is initially empty. If the total scattering of electrons and holes are not balanced then $\Delta j_i(E) \neq 0$. For simplicity we assume that the self-energy $\tilde{\Sigma}_i(E)$ is determined solely by the electron and hole densities $n(x_i; E)$ and $p(x_i; E)$ so that

$$\Sigma_i^{>}(E) = -\frac{i}{\tau_i^{>}(E)} = A_i(E)p(x_i; E), \tag{4.35}$$

$$\Sigma_i^{<}(E) = \frac{i}{\tau_i^{<}(E)} = B_i(E)n(x_i; E), \tag{4.36}$$

where $A_i(E)$ and $B_i(E)$ have to be determined according to the scattering mechanism. These forms of the self-energies are consistent with the relation $(\tilde{\Sigma}^{>})^{\dagger} = -\tilde{\Sigma}^{>}$ and $(\tilde{\Sigma}^{<})^{\dagger} = -\tilde{\Sigma}^{<}$. The retarded self-energy, on the other hand, is found as

$$\Sigma_i^R(E) = -\frac{i}{\tau_i^R(E)} + \sigma_i(E), \tag{4.37}$$

where

$$\frac{1}{\tau_i^R(E)} = \frac{1}{2}\left[\frac{1}{\tau_i^{<}(E)} + \frac{1}{\tau_i^{>}(E)}\right], \tag{4.38}$$

and $\sigma(E)$ is related to $\tau_i^R(E)$ via a Kramers-Kronig relation[167]

$$\sigma_i^R(E) = \frac{1}{2\pi}\mathrm{P}\int \frac{dE'}{(E - E')\tau_i^R(E)}, \tag{4.39}$$

P denoting the principal value of the integral. The self-consistent solution is obtained by solving the densities $n(x; E)$ and $p(x; E)$, and the self-energies $\Sigma_i^{>}(E)$ and $\Sigma_i^{<}(E)$ simultaneously.

4.2 Elastic and Inelastic Scattering

In this section a detailed solution for Equations 4.28 and 4.31 is presented for a single point scatterer in order to demonstrate the use of nonequilibrium Green's

functions. Let us consider the case of a single scatterer at $x = x_o$ with self-energy $\tilde{\Sigma}(E)$. Then, Equation 4.28 becomes

$$\begin{aligned} n(x;E) = n_o(x;E) \;&+\; 2\,\mathrm{Im}\,[G_o^R(x,x_o;E)\{\Sigma_*^R\}G_o^<(x_o;x:E)] \\ &-\; i\,[G_o^R(x,x_o;E)\{\Sigma_*^<\}G_o^A(x_o,x:E)]. \end{aligned} \tag{4.40}$$

Obviously, it is necessary to calculate the noninteracting Green's functions first. We assume that the noninteracting system is uniform, i.e., a 1D free EG. Thus, one represents the noninteracting Green's functions in terms of the occupation of the plane waves

$$G_o^<(x,x';E) = i\int dk e^{ik(x-x')} n_k \delta(E(k)-E), \tag{4.41}$$

$$G_o^>(x,x';E) = -i\int dk e^{ik(x-x')} (1-n_k)\delta(E(k)-E), \tag{4.42}$$

where n_k is the occupation number of the state described by the wave vector k and its energy is $E(k) = k^2/2m$. On the other hand, the retarded Green's function is independent of the occupation numbers and is given by

$$G_o^R(x,x';E) = \lim_{\eta\to 0^+} \frac{1}{2\pi}\int dk \frac{e^{ik(x-x')}}{E-E(k)+i\eta}, \tag{4.43}$$

Now, it is in order to specify the boundary conditions for $\tilde{\mathbf{G}}_o(E)$. For the time being we assume that the 1D device is connected to reservoirs at $x = \pm\infty$, with electrochemical potentials μ_L and μ_R for the left- and right-hand side one, respectively. These reservoirs act as current sources for the noninteracting system. However, this does not mean that we fixed the current probes; this is a point which is clarified in the following section. The current sources for the noninteracting system satisfy the blackbody boundary conditions. That is, at zero temperature the right- and left-going states with energies below μ_L and μ_R, respectively, are filled and the rest of the states are empty as given by

$$\begin{aligned} n_{k>0}(E) &= 1, \qquad E < \mu_L, \\ &= 0, \qquad E > \mu_L, \end{aligned} \tag{4.44}$$

and

$$\begin{aligned} n_{k<0}(E) &= 1, & E < \mu_R, \\ &= 0, & E > \mu_R. \end{aligned} \tag{4.45}$$

Assuming that the current flows from the left-hand side reservoir to the other one (i.e., $\mu_L > \mu_R$), the electron Green's function takes the form

$$\begin{aligned} G_o^<(x,x';E) &= \frac{2im}{k}\cos[2k(x-x')], & E < \mu_R, \\ &= \frac{im}{k}\exp[ik(x-x')], & \mu_R < E < \mu_L, \\ &= 0, & \mu_L < E, \end{aligned} \tag{4.46}$$

where $k = \sqrt{2mE}$ is the wave vector. For the retarded Green's function one has

$$G_o^R(x,x';E) = -\frac{im}{k}\exp[ik|x-x'|]. \tag{4.47}$$

The electron density and current are calculated by using $G_o^<(E)$ to yield

$$n_o(x;E) = \frac{m}{k}[\theta(\mu_L - E) + \theta(\mu_R - E)], \tag{4.48}$$

$$j_o(x;E) = [\theta(\mu_L - E) - \theta(\mu_R - E)], \tag{4.49}$$

Note that, m/k is the 1D density of states, so that the density of electron states is given by the density of states times the electron occupancy. The energy current density, on the other hand, is related to the excess density of right-going electrons as compared to the left-going ones.

Determining the Green's functions and the related entities for the case of noninteracting uniform system, we proceed to solve the problems concerning single elastic and inelastic scatterers. We are not concentrating on a specific scattering mechanism, so we carry out a parametric analysis of scattering[‡]. To this end, the irreducible self-energy is written as it is done in Section 4.1, Equation 4.35-4.37

$$\Sigma^>(E) = -\frac{i}{\tau^>(E)}, \tag{4.50}$$

[‡]Note that, for the multisubband case such parametric analysis would be ambiguous, since one has to determine a number of intra- and intersubband scattering times τ_{nm}.

$$\Sigma^{<}(E) = \frac{i}{\tau^{<}(E)}, \tag{4.51}$$

$$\Sigma^{R}(E) = -\frac{i}{\tau^{R}(E)} + \sigma(E). \tag{4.52}$$

For the time being we assume that $\tau^{<} = \tau^{>} = \tau^{R} = \tau$, and suppress the dependence on energy E. For the calculation of Green's functions we take $\mu_L > E > \mu_R$ so that there is a steady current flow from the left-hand side reservoir to the other one, i.e., $j_o = 1$. Following the procedure described in Section 4.1 one finds

$$n(x;E) = \frac{m}{k} + 2\left(\frac{m}{k}\right)^2 \mathrm{Im}\left\{e^{ik|x-x_o|}e^{ik(x_o-x)}\frac{\left(\sigma - \frac{i}{\tau}\right) - i\frac{m}{k}\left(\sigma^2 + \frac{1}{\tau^2}\right)}{\left(1+\frac{m}{k\tau}\right)^2 + \left(\frac{m\sigma}{k}\right)^2}\right\} + \left(\frac{m}{k}\right)^2 \frac{\frac{1}{\tau} + \frac{m}{k}\left(\sigma^2 + \frac{1}{\tau^2}\right)}{\left(1+\frac{m}{k\tau}\right)^2 + \left(\frac{m\sigma}{k}\right)^2}, \tag{4.53}$$

for the density of electrons and

$$j(x;E) = \frac{1+\frac{m}{k\tau}}{\left(1+\frac{m}{k\tau}\right)^2 + \left(\frac{m\sigma}{k}\right)^2}, \tag{4.54}$$

for the current (or equivalently the transmission probability $T = |t|^2$ for the scatterer). It can be shown that the density $n(x;E)$ is constant and equal to Tm/k for $x > x_o$ and oscillates for $x < x_o$ with a period π/k.

An elastic scatterer is defined by the condition $\tau \to \infty$, that is the carriers are not absorbed and/or injected by the scatterer. For this condition being satisfied the transmission probability becomes

$$T = \frac{1}{1+\left(\frac{m\sigma}{k}\right)^2}, \tag{4.55}$$

which can be found also by solving the 1D Schrödinger equation for a δ-function potential of strength σ. Using Equation 4.53, the density of electrons is expressed as

$$\begin{aligned} n(x;E) &= \frac{m}{k}\left[1 + R + 2R^{1/2}\sin\{2k(x_o - x) - \varphi\}\right], && x < x_o, \\ &= \frac{m}{k}T, && x > x_o, \end{aligned} \tag{4.56}$$

where $R = 1-T$ is the reflection probability of the scatterer and $\varphi = \sin^{-1} R^{1/2} = \cos^{-1} T^{1/2}$ is the phase associated with the scatterer. The rederivation of these well-known[106,168] results by using the present formalism shows that it is possible to include any kind of quantum mechanical scatterer in the transport system under consideration, which is a main advantage of the present approach. That is, one does not have to calculate the equilibrium Green's function in the presence of all the elastic scatterers and include the effect of the inelastic scattering afterwards.

An inelastic scatterer absorbs and emits the carriers, so that $\tau < \infty$. As a simple example we take $\sigma = 0$ and only inelastic scattering would take place. Equation 4.53 gives

$$T = \frac{1}{1 + \frac{m}{k\tau}}, \tag{4.57}$$

for the transmission probability and

$$\begin{aligned} n(x;E) &= \frac{m}{k}\left[1 + R - 2R\,\sin\{2k(x_o - x)\}\right], && x < x_o, \\ &= \frac{m}{k}T, && x > x_o, \end{aligned} \tag{4.58}$$

for the electron density. Note that, for the inelastic scatterer the oscillations for $x < x_o$ have the amplitude R as compared to $R^{1/2}$ for the elastic scatterer. Similarly, for the inelastic one there is no phase associated with the scattering. This is due to the different scattering mechanism for two types of scatterers. The absorption and emission of carriers are not coherent processes, thus the probability amplitudes such as $R^{1/2}\exp(i\varphi)$ does not appear in the results for inelastic scattering. Nevertheless, the charge density is still oscillating for $x < x_o$. Therefore, the model inelastic scatterers can not be visualized by completely phase randomizing reservoirs. Such a phase randomization may be possible by using current probes described in Section 4.3 only. Clearly, the solutions for the inelastic scatterer can not be obtained by making use of the Schrödinger equation, since the Hamiltonian is not Hermitian and thinking only in terms of electrons is not sufficient.

For the systems analyzed above, the current is constant and equal to T throughout the device. This is due to the fact that the electron and hole densities

and their interactions with the scatterer are completely symmetric. A current nonconserving scatterer has the property $\tau^< \neq \tau^>$. For $\sigma = 0$, one finds the transmission probabilities for electrons and holes as

$$T_e = \frac{1 + \frac{m}{k\tau^<}}{\left(1 + \frac{m}{k\tau^R}\right)^2}, \tag{4.59}$$

$$T_h = \frac{1 + \frac{m}{k\tau^>}}{\left(1 + \frac{m}{k\tau^R}\right)^2}, \tag{4.60}$$

respectively, where $\tau^R = 2\tau^<\tau^>/(\tau^< + \tau^>)$. The discontinuity in the current is found as

$$\Delta j = \frac{\frac{m}{k\tau^>} - \frac{m}{k\tau^<}}{\left(1 + \frac{m}{k\tau^R}\right)^2}. \tag{4.61}$$

Apparently, the current is drawn out of (emitted into) the circuit by the inelastic scatterer if $\tau^> < \tau^>$ ($\tau^> > \tau^>$), that is if the electrons (holes) interact with the scatterer more strongly. In the following section a model current probe is built using these current nonconserving scatterers.

Finally, we want to demonstrate how inelastic scattering prevents the phase coherence. It is well known that two identical barriers $\lambda/2$ apart yield resonant tunneling. That is, when the energy of the incident state matches the energy of the quasi-bound state (that formed in the well between the barriers), the transmission probability becomes unity. Clearly, the particle has to retain its phase between the barriers in order to have this resonance effect. In Figure 4.1 transmission probability for such a structure is shown in the presence of inelastic scattering in both the barriers and well. As depicted, the resonance feature diminishes with increasing inelastic scattering rate. It is appropriate to define an inelastic scattering length L_{in}, which is equal to the average length over which the particle moves without suffering an inelastic scattering event. This length is calculated for the uniform system as[‡]

$$L_{in} = \frac{v\tau}{n_{sc}}, \tag{4.62}$$

[‡]Here and in the rest of the chapter, when referring the numerical values of τ, the self-consistent corrections are not taken into account. In other words, we use the τ value that would be obtained for $n = m/k$ only, for the sake of simplicity. The scale of length is such that the wavelength of the electrons is $\lambda = 2\pi/k$.

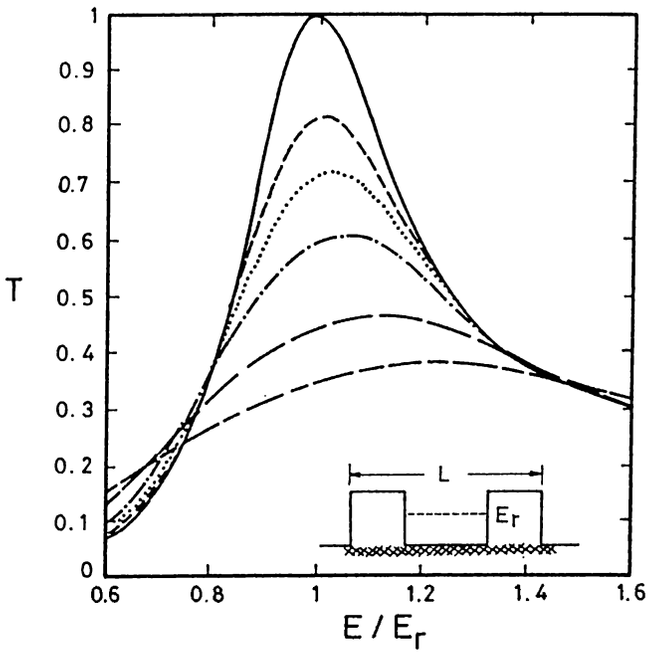

Figure 4.1: Resonant tunneling in the presence of inelastic scattering
The transmission probability as a function of energy. E_r is the resonance energy. The curves correspond to $L_{in}/L = \infty$, 60, 30, 15, 6, and 3, in which order the peak current decreases. The length of the device is L including the barriers and well. Inelastic scattering is present throughout the device.

where $v = k/m$ is the velocity of electrons and n_{sc} is the density of scatterers. The peak in the transmission probability curve disappears for $L_{in} \gtrsim L$, where L is the length of the device. These results are the generalization of the earlier analysis of the effects phase coherence on resonance effects by Büttiker.[169,170]

4.3 Voltage and Current Probes

For voltage measurements, we follow the reasoning of Payne.[76] That is, the voltage measured in a transport measurement is an electrochemical potential. Therefore, in order to measure a voltage it is necessary to define electrochemical potentials for a nonequilibrium system, despite the fact that this does not sound reasonable. However, if one recalls that the electrochemical potentials of the

fictitious reservoirs at $x \rightarrow \pm\infty$ are well defined quantities μ_L and μ_R, and these reservoirs satisfy blackbody boundary conditions, it becomes easier to understand the action of a voltage probe. In the absence of any scatterers all the right- and left-going states are occupied up to the electrochemical potential of the corresponding reservoirs. By inserting a barrier on the device, some of the right-going states will be reflected from the barrier and thus part of the left-going carriers will be originated from the left-hand side reservoir. Now consider a non-invasive voltage probe between the left-hand side reservoir and barrier. The number of electrons scattered into the voltage probe will have a larger contribution from the left-hand side reservoir than the other one, due to the reflection at the barrier. Therefore, the voltage measured by the probe is closer to μ_L than μ_R. In the case of unity reflection this voltage is equal to μ_L, since there is no current and the probe is short-circuited to the left-hand side reservoir. In the highly transparent barrier limit, on the other hand, the contribution of the left- and right-hand side reservoirs are equal and thus the voltage measured is halfway between μ_L and μ_R. Clearly, in order to simulate the operation of a voltage probe we have to find a way of differentiating the electrons originated from left- and right-hand side reservoirs. The above intuitive argumentation is beneficial for visualizing the operation of a voltage probe. However, it is not consistent with the present nonequilibrium approach, since we consider only the electrons and not the holes. Nevertheless, for $\mu_L > E > \mu_R$ there are right-going electrons and left-going holes in the absence of the barrier. Therefore, it is possible to call right- and left-going states as electrons and holes, respectively, in order to meet this requirement. This, in addition, solves the problem of differentiating the carriers originated from different reservoirs.

The voltage measured by the probe μ_p can be defined by exploiting the conservation of current argument. Assume that an ideal wire attaches the probe (in the simplest case is an inelastic scatterer) to a reservoir in thermodynamic equilibrium with electrochemical potential μ_p. The left- and right-hand side reservoirs have excess electrons and holes, respectively, as compared to the reservoir of the voltage probe. At the position of attachment of the probe to

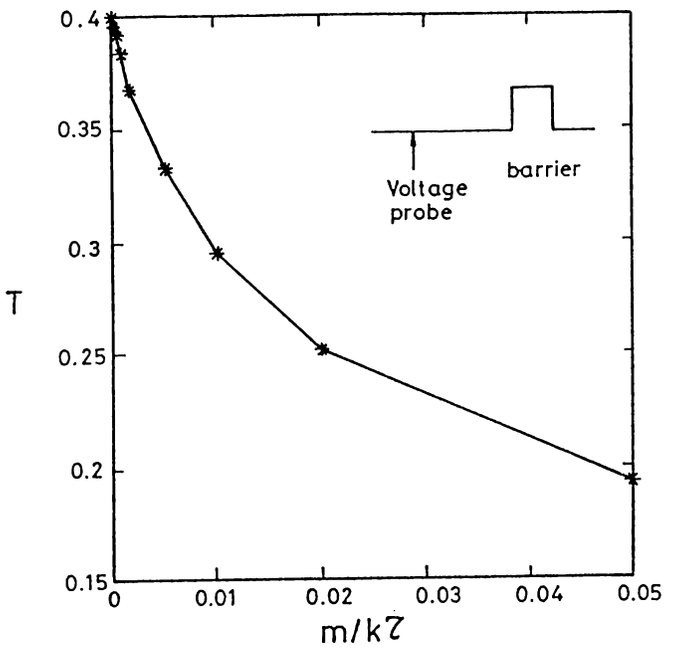

Figure 4.2: Effects of a voltage probe on the device characteristics
The device is an elastic barrier extending from $x = -1$ to 1. The transmission probability is $T = 0.4$. The voltage probe is a single inelastic scatterer located at $x_p = -4.75$. The transmission probability is shown as a function of the inelastic scattering rate $m/k\tau$.

the device x_p, there are $n(x_p; E) \times (\mu_L - \mu_p)$ electrons and $p(x_p; E) \times (\mu_R - \mu_p)$ holes which cause a flow into the probe reservoir. The net current going into this reservoir has to be zero, however, since it is a voltage probe. Therefore, the flow of electrons and holes has to be balanced. That is, the electrochemical potential of the probe reservoir has to be

$$\mu_p = \frac{n(x_p; E)\mu_L + p(x_p; E)\mu_R}{n(x_p; E) + p(x_p; E)}, \tag{4.63}$$

where we assumed that the propagation of the electrons and holes through the ideal wire, which connects the scatterer to the reservoir, does not differ. The difference of the electrochemical potentials is taken to be small, so that it is possible to take n and p constant throughout the energy range $[\mu_R, \mu_L]$. Equation 4.63 is reminiscent of the counting argument by Landauer.[23] A similar

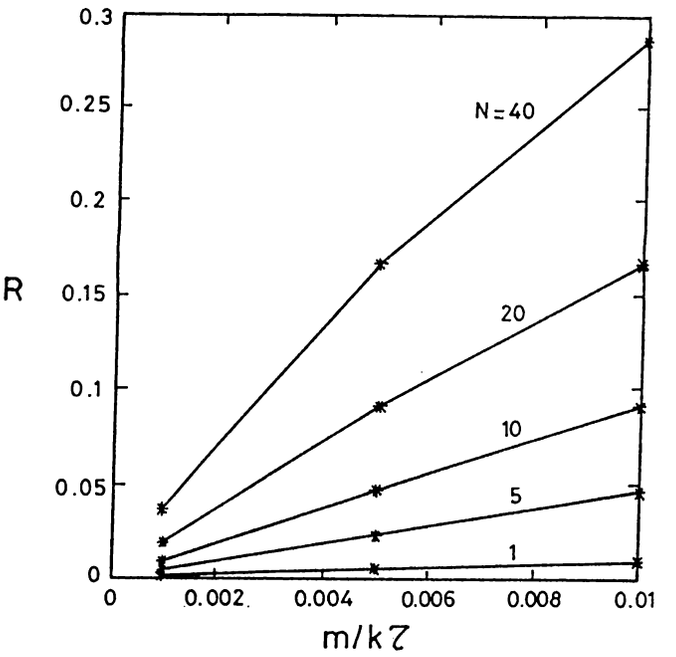

Figure 4.3: Reflection from extended voltage probes
The reflection probability is shown as a function of the inelastic scattering rate $m/k\tau$. The scatterers are placed $\Delta x_p = 1$ apart. The curves correspond to 1, 5, 10, 20, 40 scatterers in which order reflection increases.

expression was found by Büttiker[51] making use of the multilead form of the Landauer formula.[50]

One of the important questions concerning the characteristics of a voltage probe is its effect on the device. A *non-invasive* probe is the one which has an infinitesimal influence on the circuit. Such a voltage probe can be obtained by use of, for example, an STM.[171,172] On the other hand, for the electron waveguide structures the leads attached to the voltage measurement devices have similar dimensions as the waveguide itself, thus yield invasive measurements.[51] In the present formalism the coupling of the voltage probe to the device is determined by the self-energy due to the inelastic scattering caused by the probe. Therefore, the probe can be made non-invasive in the limit $\tau \to \infty$. In Figure 4.2 the performance of such a voltage probe is shown. We find that the voltage measured by the probe stays fairly constant as τ decreases. However, the increasing rate

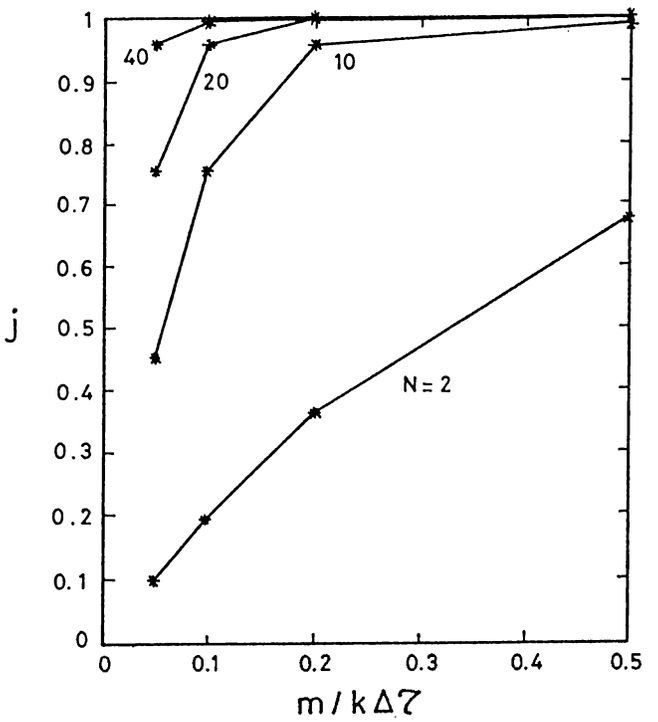

Figure 4.4: Current injection capability of extended current probes
The fraction of current that is injected by the current probe is shown as a function of the difference of the inelastic scattering rates $m/k\Delta\tau$. The curves correspond to 2, 10, 20, 40 scatterers in which order fraction increases.

of scattering by the probe leads to a decrease in the current passing through the device. We conclude that a non-invasive probe has to have $k\tau/m < 10^{-2}$, so that the percent change in current is $< 1\%$. In Figure 4.3 the characteristics of extended voltage probes are shown. Expectedly, as the number of scatterers increases the current through the circuit decreases. The above rule of thumb for the non-invasive probes now takes the form $k\tau/m < 10^{-2}/N_{sc}$, where N_{sc} is the number of scatterers in the probe. For probes with more than one scatterer the electrochemical potential of the reservoirs connected to each of these scatterers is different. Therefore, it is necessary to include an averaging over the probe as well. In the next section we demonstrate the use of voltage probes in detail.

Next we build a current probe by using inelastic scatterers. The current and voltage probes are quite different in their operation. It is possible for a voltage probe to decrease the amount of interaction with the device in such

a way that the measuring action is still present. A current probe, however, has to have a substantial amount of interaction with the device since it has to draw or inject all the current through the device. Therefore, a non-invasive current probe may be defined as such that all the incoming current is absorbed and no reflection takes part. Examining Equation 4.34 one observes that the discontinuity in current for a single current nonconserving scatterer goes as $1/\tau$ in the lowest order. The reflection probability, on the other hand, has the same dependence on τ^R. Therefore, using a single scatterer it is not possible to obtain a non-invasive current probe. In order to obtain zero reflection, we exploit the resonance structure for the current probes. Mention of resonance of two identical barriers $\lambda/2$ apart is made before. In resonance they do not give rise to reflection in steady-state as a result of the constructive interference of the forward-going waves. In fact, the resonance condition for current nonconserving inelastic scatterers is different since the current is discontinuous at the position of the scatterers. Nevertheless, a simple model by using two inelastic scatterers shows that the minimum reflection is obtained when the distance between the scatterers is very close to $\lambda/2$. Therefore, we do not use the exact expression for the resonance condition. It can be shown that a non-invasive current probe can not be implemented by using only two scatterers, neither. As mentioned above such a probe has to draw all the current passing through the circuit without any reflection. In the two-scatterers case, in order to increase the amount of current drawn by the probe one has to decrease $\tau^<$ which causes τ^R to decrease and reflection to increase. Thus, it is necessary to put enough number of weak current nonconserving inelastic scatterers $\lambda/2$ apart so that the resonance condition is approximately valid and the current is completely drawn out of the scatterers. In Figure 4.4 the performance of such extended current probes are shown. For all the configurations less than 1% of the current is reflected back, the rest is either drawn by the probe or transmitted through. Therefore, we conclude that the limiting quantity for a current probe is its capability of drawing current out of the system. This can be achieved by using either a few small $\tau^>$ scatterers or a large number of large $\tau^>$ scatterers. Since the number of scatterers determines

the size of the computations, it is more appropriate to choose the first case for our further applications. Note that, the fictitious reservoirs at $x = \pm\infty$ become irrelevant when a model current probe is present. That is, they do not draw (inject) any current out of (into) the system.

Finally, we wish to discuss the relevance of the probe models developed in this section to the real devices. In fact, the experiments are being carried out in multilead geometries. That is, the connections to the probes are similar to the device under consideration and not ideal wires. Nevertheless, for example, the single scatterer voltage probe may be used to simulate an STM voltage probe. An extended voltage probe, on the other hand, is reminiscent of the *phase-averaged* voltage probe introduced by Büttiker.[173] The current probes defined above are less realistic structures. For instance, a current probe has the non-invasive character only for a certain energy or electron density. Varying either of these, the separation between the scatterers will deviate from $\lambda/2$ so that appreciable reflection from the probe will take place. The real current probes, on the other hand, are wide reservoirs similar to those discussed in Chapter 2. The dilution effect[77] prevents the reflection from the probe into the device and the current is drawn out of the system away from the device, in this reservoir. However, it is still possible to extract some information about the effect of probes on measurements by using these *toy models* for probes. A more realistic investigation of multiprobe and multilead devices is beyond the scope of this study.

At this point it is in order to survey the existing literature on nonequilibrium quantum transport and make a comparison with the present method. Although there are numerous studies on quantum transport in mesoscopic systems, the attention has recently been focused on the effects of inelastic scattering. It is worthwhile to summarize four very recent studies[167,174–176] which present different approaches to the problem of quantum transport. As mentioned before, Datta[167] used the Keldysh formalism to derive a quantum kinetic equation. Assuming local thermodynamic equilibrium he obtained a generalization of Landauer formula for a continuous distribution of probes. The weakness of his method is the heuristic inclusion of external current in the system. This way, the quantum mechanical

interaction between the current source and device is lost. This problem is avoided in the present approach by using the noninteracting Green's function $\tilde{G}_o$ obtained for a nonequilibrium state. Later D'Amato and Pastawski[174] used the multiprobe Landauer formula[51] to study the conductance of linear chains with inelastic scatterers. That study and the present approach uses the fact that the voltage probes does not draw any current out of the device. This is in agreement with the proposition[50,51] that a voltage probe and an inelastic scatterer are of the same nature and one may be used in place of the other. However, there is an important difference between the two methods. As discussed in the previous section the current conservation in the present approach is obtained by varying the strength of the inelastic scatterer. D'Amato and Pastawski,[174] on the other hand, varied the *electrochemical potential* of the reservoirs along the chain. That is, emphasis is placed on the inelastic scattering and voltage probe aspects in this study and Reference 174, respectively. Recently Pernas and coworkers[175] employed a Keldysh technique to find the potential variations across a constriction. In fact, their method corresponds to a closed system approach to quantum transport as described in Section 1.4, since the noninteracting Green's function belongs to two noninteracting systems without any connection in between. Very recently Feng[176] included the effect of phase breaking scattering in quantum transport by using a local phonon interaction. He derived a generalized Landauer formula for two-probe case which takes the electron-phonon interaction into account. The present formalism, in contrast to some existing ones, does not yield open expressions for conductance[175,176] or transmission probability.[167,170] Nevertheless, as shown in this chapter it is capable of solving widely spread problems including those solved by using various techniques described above. In addition, modeling of probes and use of a nonequilibrium state as the noninteracting state provide unique advantages to the present approach.

4.4 Multiprobe Measurements

We mentioned in Chapter 1 that in the mesoscopic regime the means of measurement are important and may interfere with the ideal results, which are obtained by a fully non-invasive measurement. This last sentence, actually, is a restatement of the well known quantum mechanical principle about measurement. In this section we exemplify the effects of the measurement device on the system under consideration by using the probes developed in the preceding section. This question, in fact, was first pointed out in the work by Engquist and Anderson,[26] and later emphasized by Büttiker[50,51] and recently by Levinson.[106]

First we focus our attention on the non-invasive measurement method. It has long been discussed that[75,77] it is possible to prevent the interference of the current probe and the device by adiabatically injecting the current. That is, the current probe is put on a wider portion of the device and then the channel is narrowed down to the size of the device, this tapering being smooth on the length scale of the electron wavelength. It has been shown[81,89,106] that such a geometry yields adiabatic evolution of states and reflection from the ends of the device can be neglected to a good approximation. In the model presented in this chapter, the fictitious reservoirs at $x \to \pm\infty$ may be visualized as current probes which are connected adiabatically to the device. Since we assumed that they satisfy the blackbody boundary conditions, these probes are non-invasive. In the preceding section it is shown that a non-invasive voltage probe can be obtained by decreasing the interaction between the device and the single inelastic scatterer, that is, by increasing τ. Thus, in the limit $\tau \to \infty$ the presence of the probe does not have any effect on the device and the voltage given by Equation 4.63 is measured non-invasively. Consequently, a non-invasive measurement is the one obtained without using any *real* probes. In Figure 4.5 we present the results of non-invasive measurement for the device that is investigated throughout this section. It consists of a potential barrier extending from $x = -1$ to $x = 1$, with transmission probability $T = 0.4$ in the absence of any other interactions. The energy is so chosen that the wave vector k is unity. The device, in turn covers

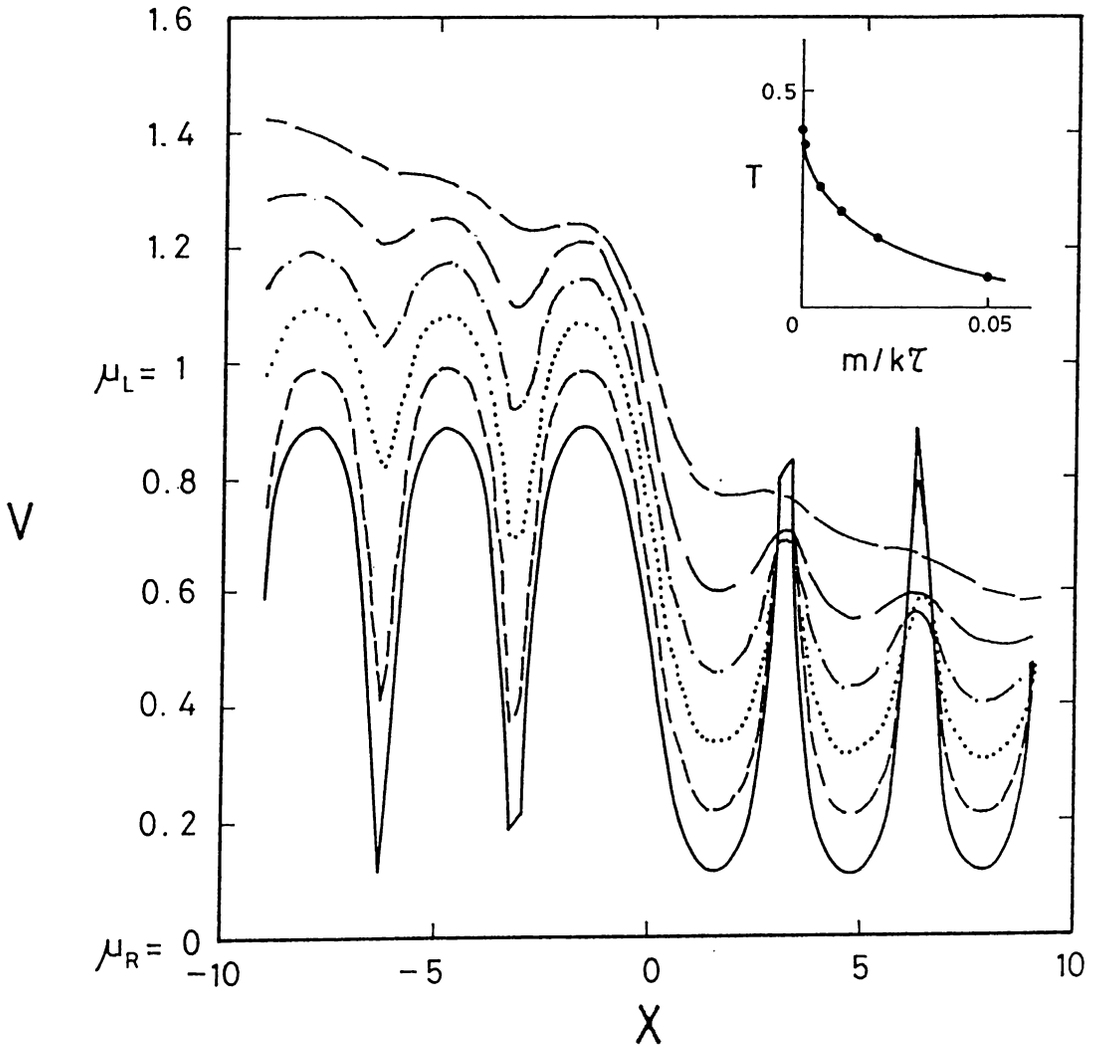

Figure 4.5: Potential oscillations near a barrier
The voltage at the position of the inelastic scatterers calculated for the device described in the text. The full, dashed, dotted, dash-dotted, long-dashed, and long-short-dashed curves correspond to $L_{in} = \infty$, 250, 50, 25, 12.5, and 5, respectively. The length of the device is $L = 18$. The curves are offset by 0.1 for clarity. The inset shows the transmission probability through the device as a function of the inelastic scattering rate.

the range from $x = -9$ to $x = 9$, and consists of the barrier and, in addition, inelastic scattering centers separated by $\Delta x = 0.25$. the strength of the inelastic scattering is varied so that the inelastic scattering length varies between $L_{in} = \infty$

and $L_{in}5$. Apparently, as the strength of the inelastic scattering increases, the transmission probability through the device decreases. This is a result of the additional resistance caused by the inelastic scattering. In fact, one can show that for a barrier consisting of solely inelastic scatterers the transmission probability is approximately proportional to L_{in}/L, which yields the Ohm's law for electrical resistance. In contrast, for an elastic barrier T decreases exponentially with length.

The most important observation on Figure 4.5 is the oscillation of the electrochemical potential along the channel. This point has been investigated earlier[26,106] and an extensive study by Büttiker[168] shed light on the importance of these oscillations. As studied in Section 4.2, the incident and reflected waves interfere with each other to form an oscillating electron density. This, in turn, gives rise to oscillation of the local electrochemical potential as defined by Equation 4.63. This is the voltage that would be measured by a non-invasive voltage probe. Büttiker[168] defined *phase-sensitive*, *phase-averaged* and *phase-insensitive* voltage probes. The non-invasive probe we are using has the counterpart phase-sensitive probe in his nomenclature. An extended voltage probe which described in the preceding section, which still has non-invasive measuring action, is a phase-averaged probe. Büttiker himself pointed out the conceptual difficulties associated with a phase-insensitive probe. Engquist and Anderson[26] and Levinson[106] earlier discussed that the thermal averaging washes out the oscillations. However, it is not clear how this averaging does not yield the phase-averaged result of Büttiker, but Landauer's original result[22,23] instead. The answer is that, their *averaging* procedure is just to take out the terms with nonzero phases. Therefore, in the present context it is not possible to obtain the Landauer's result $G \sim T/R$ by a rigorous derivation. Note that, we ignored the self-consistent potential associated with the nonequilibrium carriers. As mention made in Section 1.4, there is a controversy about the relevance and importance of this self-consistent potential. Nevertheless, recently Büttiker[79] showed that such a treatment, which in turn ignores the potential oscillations, yields the acquired answer.

As depicted in Figure 4.5, the period of potential oscillations is π, i.e., $\lambda/2$, and its amplitude decays away from the barrier when $L_{in} < \infty$. These features are reminiscent of Friedel oscillations in charge density near the surface of a metal. Clearly, the oscillations die out a few L_{in} away from the barrier. This leads us to an important (although well known) conclusion that if the distance between the voltage probes and the structure to be investigated is more than a few inelastic scattering length, then the measurement is a two-probe measurement, irrespective of the geometry. For example, for $L_{in} = 5$ in Figure 4.5, the potential drop across the device is close to $\mu_L - \mu_R$.

Next, we analyze the effect of an invasive voltage measurement on the device itself. In Figure 4.6 our results are shown. We set $L_{in} = 50$ and change the position and strength of inelastic scattering for a single scatterer voltage probe. Note that, even for the weakest voltage probe the inelastic scattering rate is 20 times shorter than that of the resistive part of the device. Therefore, in principle it causes appreciable additional scattering which interferes with the unperturbed response of the device in a complicated way. In addition the current passing through the device deviates from its original value due to this additional scattering. One observation is that when the probe is located near a maximum of the electron density it has drastic effects on the electron density as well, throughout the device. Consequently, one concludes that it is possible to have important interference between the voltage probes even if they are located on the opposite sides of the barrier. On the other hand, when it is near a minimum of the electron density its effects on both current and voltage are not that dramatic. One can show that for shorter L_{in}, the oscillations in voltage and electron density due to the presence of the probe are suppressed to a large extent.

Lastly, we focus our attention on the effects of the current probes. The results are shown in Figure 4.7. Since almost whole of the current is supplied by the current source, we investigate the effect of its position and the parameter L/L_{in} on the device. We find that for long L_{in}, or in other words if the current probe is closer to the device than a few inelastic scattering length, the current differs from the original one appreciably. In addition one observes additional oscillations in

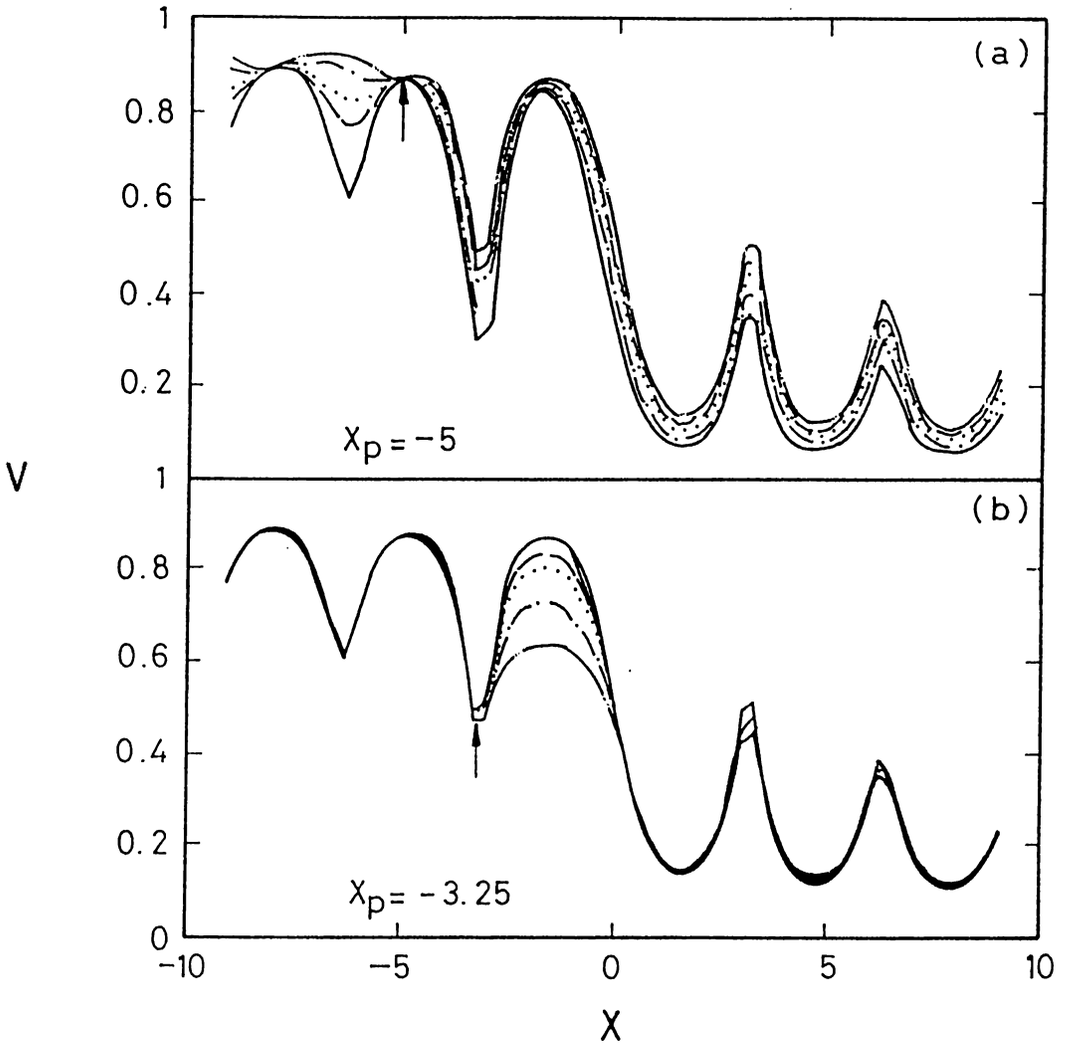

Figure 4.6: Invasive voltage measurements
The voltage at the position of the inelastic scatterers calculated for the device described in the text. The full, dashed, dotted, dash-dotted, and long-dashed curves correspond to $m/k\tau$ = 0.005, 0.1, 0.2, 0.5 and 1 for the probe, respectively. The length of the device is $L = 18$ and $L_{in} = 50$. The voltage probe is located at (a) $x_p = -5$, near a maximum of the electron density and (b) $x_p = -3.25$, near a minimum of the electron density. The transmission probability is T = 0.31, 0.27, 0.25, 0.20 in which order τ decreases in (a); and is almost constant and equal to $T = 0.31$ in (b).

voltage caused by the probe. On the other hand, if the probe is further away (in our case if L_{in} is smaller), its effect on the current is unimportant, still affecting

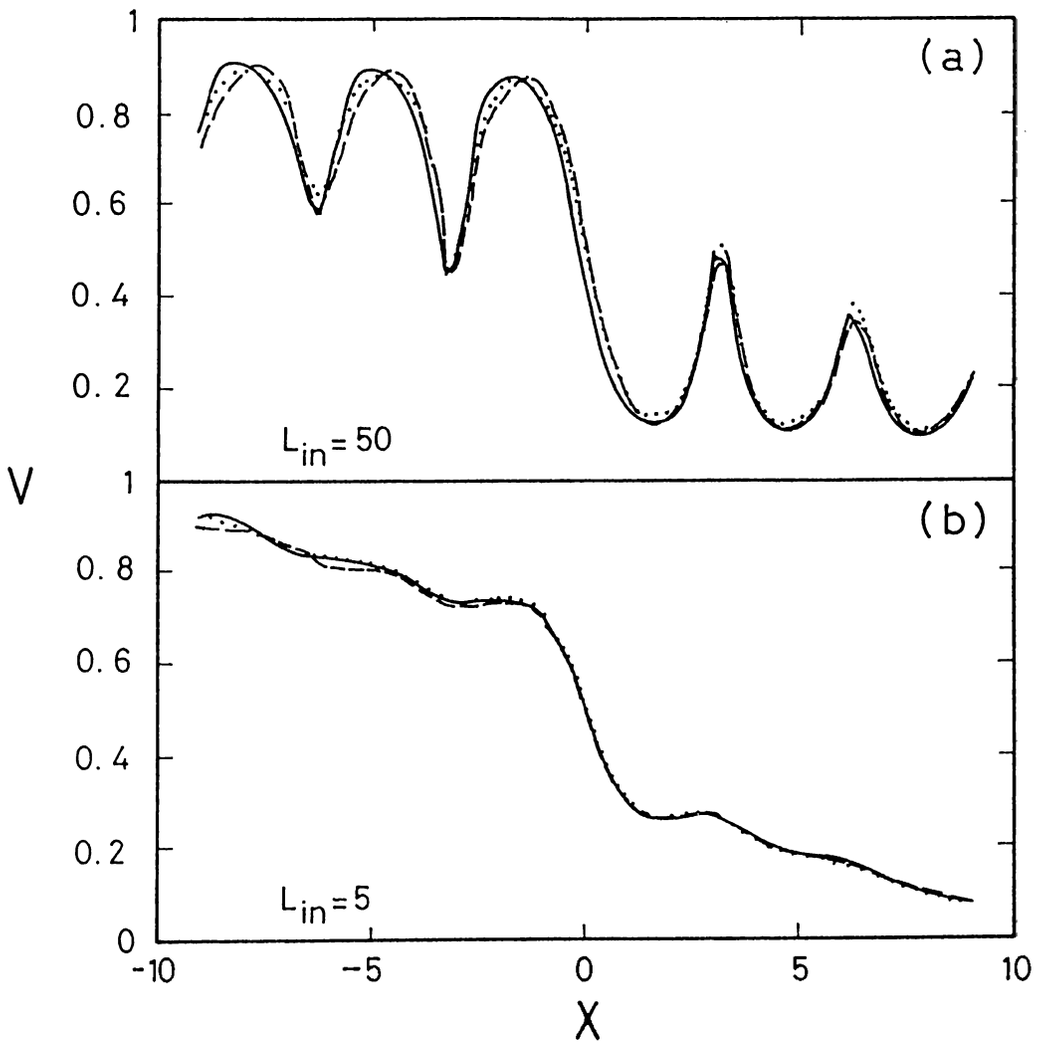

Figure 4.7: Effects of the current probe on the device characteristics
The voltage at the position of the inelastic scatterers calculated for the device described in the text. The full, dashed, and dotted, curves correspond to adiabatic injection, and a current probe at $x = -10$ and -12, respectively. The current probe consists of 10 scatterers with $m/k\Delta\tau = 0.5$. The length of the device is $L = 18$ and (a) $L_{in} = 50$ and (b) $L_{in} = 5$. The transmission probability is $T = 0.31$, 0.27, 0.27, for the respective curves in (a); and is almost constant and equal to $T = 0.15$ in (b).

the voltage to a small extent.

At this point it is in order to comment on the findings of this section. First we wish to recall that the system under consideration is assumed to have a

single occupied subband and the screening length is small compared to the Fermi wavelength and other relevant length scales. We find that a scatterer give rise to density and potential oscillations which decay exponentially with decay length L_{in} going away from the scatterer. These oscillations may be measured by a non-invasive voltage probe, such as an STM. An extended probe, on the other hand, washes out these oscillations. Nevertheless, the resulting conductance is not the one conjectured by Landauer,[22,23] but a phase-averaged one as calculated by Büttiker.[168] Therefore, for non-invasive measurements the multiprobe Landauer formula[50,51] is the relevant one. If the invasive voltage probes are closer to the scatterer than an inelastic scattering length, they influence the characteristics of the device. Otherwise, even though they affect the measured values one has an effectively two-probe measurement. The current probes, similarly may interfere with the device if they are close to the device. In the general case all the probes are coupled to both the device and each other.

4.A Green's Function Technique for Nonequilibrium Processes

In this appendix the Keldysh technique for nonequilibrium Green's functions is summarized. The related references and detailed information can be found in References 163 and 164. In Reference 162 a similar technique was developed and applications of it were presented.

Consider a nonequilibrium system described by noninteracting Hamiltonian $\hat{H}_o$ and the interaction $\hat{H}_i$. The noninteracting field operators $\hat{\psi}_o$ can be defined in terms of the eigenstates of $\hat{H}_o$ and the occupation number of these states. They satisfy the usual equal-time anti-commutation relation

$$\{\hat{\psi}_o^\dagger(\mathbf{r},t),\hat{\psi}_o(\mathbf{r}',t)\} = \delta(\mathbf{r}-\mathbf{r}'), \tag{4.64}$$

and equation of motion

$$i\frac{\partial\hat{\psi}_o}{\partial t} = [\hat{\psi}_o,\hat{H}_o], \tag{4.65}$$

in the interaction representation. It is more convenient to describe the interacting system by using the S-matrix which is defined as

$$\hat{S}(t,t') = \hat{T}\left\{\exp\left[-i\int_{t'}^{t} d\tau\ \hat{H}_i(\tau)\right]\right\}, \tag{4.66}$$

where $\hat{T}$ is the time ordering operator. The S-matrix satisfies the equation of motion

$$i\frac{\partial \hat{S}}{\partial t} = \hat{H}_i(t)\hat{S}(t,t'), \tag{4.67}$$

in the interaction representation.

The objective is to calculate the expectation values of products of operators in the form

$$< \hat{T}\{\hat{L}(t)\hat{M}(t')\cdots\} >, \tag{4.68}$$

where the expectation value is defined via

$$< \cdots >= \mathrm{Tr}\{\hat{\rho}\cdots\}, \tag{4.69}$$

$\hat{\rho}$ being the density operator and the trace is taken over the all available states of the system. At this point it is necessary to use the Heisenberg representation and thus a time-independent density operator. Assuming that at $t = t_\beta$ the system has a known distribution $\hat{\rho}_\beta$, it is possible to express the expectation value in Equation 4.68 in terms of the S-matrix. This can be done by using the following property of the S-matrix

$$\hat{\rho}(t) = \hat{S}(t,t_\beta)\hat{\rho}_\beta\hat{S}^\dagger(t,t_\beta), \tag{4.70}$$

and the cyclic independence of trace. The result can be written as

$$\begin{aligned} < \hat{T}\{\hat{L}(t)\hat{M}(t')\cdots\} > &= \mathrm{Tr}\{\hat{\rho}_\beta[\hat{T}\hat{S}(t_\beta,t)\hat{L}(t)\hat{S}(t,t')\hat{M}(t')\cdots\hat{S}(.,t_\beta)]\}, \\ &\equiv \mathrm{Tr}\{\hat{\rho}_\beta\hat{T}_c[\hat{S}_c\hat{L}_o(t)\hat{M}_o(t')\cdots]\}. \end{aligned} \tag{4.71}$$

where in the second line the operators are of the interaction representation and the subscript c denotes the integration contour. This contour goes from t_β to t_α (a sufficiently large time for which all the interactions turn off), and then back to t_β. The new time ordering operator $\hat{T}_c$ orders the operators on this integration

contour. This strange contour requires that the time on the forward branch (that going from t_β to t_α) and the reverse branch (that going from t_α to t_β) has to be differentiated. For example, product of two operators may have four different values, depending on the branch that the time variables belong. Denoting the forward and reverse branches by t_+ and t_-, respectively, four Green's functions can be defined for the noninteracting case.

$$G_o^<(\mathbf{r},t;\mathbf{r}',t') = -i\ \mathrm{Tr}\{\hat{\rho}_\beta \hat{T}_c[\hat{\psi}_o(\mathbf{r},t_+)\hat{\psi}_o^\dagger(\mathbf{r}',t'_-)]\}, \tag{4.72}$$

$$G_o^>(\mathbf{r},t;\mathbf{r}',t') = -i\ \mathrm{Tr}\{\hat{\rho}_\beta \hat{T}_c[\hat{\psi}_o(\mathbf{r},t_-)\hat{\psi}_o^\dagger(\mathbf{r}',t'_+)]\}, \tag{4.73}$$

$$G_o^T(\mathbf{r},t;\mathbf{r}',t') = -i\ \mathrm{Tr}\{\hat{\rho}_\beta \hat{T}_c[\hat{\psi}_o(\mathbf{r},t_+)\hat{\psi}_o^\dagger(\mathbf{r}',t'_+)]\}, \tag{4.74}$$

$$G_o^{\bar{T}}(\mathbf{r},t;\mathbf{r}',t') = -i\ \mathrm{Tr}\{\hat{\rho}_\beta \hat{T}_c[\hat{\psi}_o(\mathbf{r},t_-)\hat{\psi}_o^\dagger(\mathbf{r}',t'_-)]\}. \tag{4.75}$$

Note that, in the first two equations $\hat{T}_c$ has no real ordering effect. These are electron and hole Green's functions, respectively. The last two are, in turn, are the time ordered (or causal) and anti-time ordered Green's functions, respectively. The Green's functions for the interacting system are defined similarly (by including $\hat{S}_c$ in the time ordered products) which are equivalent to those given in Equation 4.3-4.6. It is possible to use the Feynman diagram technique to calculate the interacting Green's functions. However, attention has to be placed on the proper time ordering. In particular, both The Green's functions and self-energies can be cast in a matrix form as described in Section 4.1. The Green's function matrix satisfies a Dyson equation, given in Equation 4.1, which is formally the same as the nonequilibrium one.

Chapter 5

Summary and Conclusions

In this study mainly the ballistic transport and tunneling in laterally confined systems are investigated. Therefore, the subject matter is the geometry-specific features of electronic transport in mesoscopic devices. This features are, in fact, closely related to the solutions of Schrödinger equation. In Chapters 2 and 3 solution procedures for Schrödinger equation has been developed for quantum point contacts and scanning tunneling microscopy, respectively. Using non-invasive probes one measures the quantities obtained by using these methods. However, in reality, neither the probes are completely non-invasive nor the system at hand is free of any inelastic scatterers. Therefore, it is necessary to devise a formalism to study the quantum transport more rigorously. This is done in Chapter 4 by using nonequilibrium Green's function technique. This last chapter summarizes the contributions made by this thesis work and emphasizes the important points by way of conclusions.

Quantum Point Contacts

The quantum point contacts have been used in several experiments since their introduction by Delft-Philips collaboration[58] and Cambridge group.[59] Most of the novel applications of QPC take place in finite magnetic field. Nevertheless, it is necessary to understand the transport properties of QPC in zero-field case, in order to extract information about their structural parameters and transport

properties. This subject is thoroughly analyzed in Chapter 2.

- A conductance formula is derived in conjunction with Landauer formulae and simple quantum mechanical explanation of the quantization of conductance given in Section 1.2. Namely, the coefficients of the subband wave functions in the current carrying states determine the conductance in Equation 2.18.

- The discrepancy between the experimental results[58,59] and the results obtained for uniform constrictions, which are studied in Section 2.2, was one of the important problems at the early stages. Investigating the effects of temperature, finite bias and geometry we found that the geometrical structure of the constriction plays an essential role leading to adiabatic evolution of states through the constriction and thus causing the resonance structure to disappear.

- The solutions of Schrödinger equation as found in Chapter 2 have strictly 1D analogues. This analogy both simplifies the underlying physics and makes it possible to investigate some novel phenomena. Among these, adiabatic evolution and resonant tunneling through quasi-0D states, as studied in Section 2.3, are the most important ones. Although the adiabatic transport has been investigated experimentally, resonant tunneling demands further experiments.

- At the same time, the quasi-1D systems have some essential differences from the strictly 1D counterparts. They result from the subband structure and become important when the intercation between the subbands is appreciable. The enhanced backscattering in the presence of attractive impurities in the channel is studied in Section 2.4 and constitutes an important manifestation of this feature. This effect has been verified experimentally[122] supporting the present argumentation.

Our analysis show that the semiclassical approximation or the optical analogue for the QPC explain some of the phenomena taking place. Still, the

quantum mechanical properties are essential, especially for QPC with a small number of occupied subbands.

Scanning Tunneling Microscopy

After its invention by Binnig and Rohrer[65] the scanning tunneling microscope have had wide variety of applications, mostly in surface physics. However, the physically interesting mode of operation of STM is not the scanning, but the small tip-sample distance regime. We clarified the resemblance of STM in this regime to a QPC. Using this resemblance we developed a new approach to STM in Chapter 3, which includes the tip-sample interaction and the resulting lateral confinement of the current carrying states.

- The formalism derived for QPC is adapted to STM as described in Section 3.1. The conductance expression given in Equation 3.4 covers the independent electrode, TILS and point contact regimes of STM. In addition, our approach is both more realistic as compared to the jellium approximation and simpler than the SCF pseudopotential calculations.

- The effect of the lateral confinement increases with decreasing tip-sample distance. At the same time, the real potential barrier decreases as a result of the tip-sample interaction. These two effects together yields a slowly decreasing apparent barrier height and saturation of the tunneling current. However, transport properties near the initiation of the point contact is sensitive to the shape of the tip and is not generic. The ballistic transport after the mechanical contact, on the other hand, is dominated by the plastic deformation of the tip. These findings shed light on the experiments on transition from tunneling to point contact.[69,70]

- The tip-sample interaction and its effects on tunneling are highly dependent on the lateral position of the tip. This point was analyzed in detail by Ciraci and coworkers.[134–136] Using their results for Al tip-Al sample system we calculated the tunneling current for STM of Al. Our results are in

agreement with the experimentally observed anomalous corrugation.[71,72] This conjectures that the tip-sample interaction affects the microscopy of surfaces as well.

- Electron emission from point sources is investigated for both atomically sharp tips and QPC. We find that the effective barrier arising from the lateral confinement of the current carrying states is essential for the collimated emission from these point sources. We also show that a horn-like opening improves the emission characteristics, even though it leads to nonadiabatic transport.

Earlier studies by our group[127–129] pointed out the importance of the tip-sample interaction and its effects on tunneling. These works, however, have gone unnoticed until recently. Currently, the trend of the *STM community* has been changed to study the TILS regime and tip-sample interaction in STM. We hope that in the following years more experimental studies will be carried out on this matter and this subject will be one of the main subfields of STM.

Theory of Quantum Transport

To this point, we assume that the systems we are investigating are free of any inelastic or phase-breaking scattering. In reality one has both kind of events, which interfere with the completely phase coherent results found in Chapters 2 and 3. Using nonequilibrium Green's function technique of Keldysh we study the basics of quantum transport in Chapter 4.

- By using some simplifying assumptions for inelastic and phase-breaking scattering events, we formulate a nonequilibrium theory for steady-state quantum transport. It is both formal in the sense that it follows from the most basic equations for nonequilibrium transport, and simple enough to be used in numerical calculations. Although we employ a very simple scattering mechanism for the sake of simplicity, the method can be extended to cope with more realistic cases.

- We analyze the effects of phase-breaking interaction on resonant tunneling and potential oscillations near a barrier. The former is used as a benchmark to compare the present approach to those existing in the literature. The latter, on the other hand, is the first study of potential oscillations in the presence of phase-breaking interactions.

- We model the voltage and current probes by using inelastic scatterers. This way we find some rules of thumb for the invasive character of probes used in the measurements. Invasive measurements are investigated to study the effects of the probes on the device and the plausibility criteria for invasive and non-invasive measurements. On these grounds the related Landauer formulae are examined.

Recently the phase coherence and inelastic scattering issues in mesoscopic transport have become the focus of attention. The formalism developed in this work can be used in several related problems. It, also, may be developed to study the nonequilibrium problems such as hot carrier transport as well.

Bibliography

[1] R. W. Keyes, "Miniaturization of electronics and its limits", IBM Journal of Research and Development, **32**:24, (1988).

[2] P. Guéret, P. Buchmann, K. Daetwyler, and P. Vettiger, "Resistance of very small-area ohmic contacts on GaAs", Applied Physics Letters, **55**:1735, (1989).

[3] H. Heinrich, G. Bauer, and F. Kuchar, editors, *Physics and Technology of Submicron Devices*, Springer, Berlin, (1988).

[4] R. A. Webb and R. B. Laibowitz, editors, special issue on: *"Mesoscopic phenomena and nanolithographic technology"*, IBM Journal of Research and Development, **32**(No: 3-4), (1988).

[5] M. A. Reed and W. P. Kirk, editors, *Nanostructure Physics and Fabrication*, Academic Press, New York, (1989).

[6] P. A. Lee, R. A. Webb, and B. L. Al'tshuler, editors, *Mesoscopic Phenomena in Solids*, Elsevier, Amsterdam, (1989).

[7] Y. Aharonov and D. Bohm, "Significance of electromagnetic potentials in the quantum theory", Physical Review, **115**:485, (1959).

[8] Y. Aharonov and D. Bohm, "Further considerations on electromagnetic potentials in the quantum theory", Physical Review, **123**:1511, (1961).

[9] D. Yu. Sharvin and Yu. V. Sharvin, "Magnetic flux quantization in a cylindrical film of a normal metal", JETP Letters, **34**:272, (1981).

[10] B. L. Al'tshuler, A. G. Aronov, B. Z. Spivak, D. Yu. Sharvin, and Yu. V. Sharvin, "Observation of Aharonov-Bohm effect in hollow metal cylinders", JETP Letters, **35**:588, (1982).

[11] B. L. Al'tshuler, A. G. Aronov, and B. Z. Spivak, "The Aharonov-Bohm effect in disordered conductors", JETP Letters, **33**:94, (1981).

[12] R. A. Webb, S. Washburn, C. P. Umbach, and R. B. Laibowitz, "Observation of h/e Aharonov-Bohm oscillations in normal-metal rings", Physical Review Letters, **54**:2696, (1985).

[13] V. Chandrasekhar, M. J. Rooks, S. Wind, and D. E. Prober, "Observation of Aharonov-Bohm electron interference effects with periods h/e and $h/2e$ in individual micron-size, normal-metal rings", Physical Review Letters, **55**:1610, (1985).

[14] C. P. Umbach, C. van Haesendonck, R. B. Laibowitz, S. Washburn, and R. A. Webb, "Direct observation of ensemble averaging of the Aharonov-Bohm effect in normal-metal loops", Physical Review Letters, **56**:386, (1986).

[15] G. Bergmann, "Weak localization in thin Films: A time-of-flight experiment with conduction electrons", Physics Reports, **107**:1, (1984).

[16] S. Washburn and R. A. Webb, "Aharonov-Bohm effect in normal metal–Quantum coherence and transport", Advances in Physics, **35**:375, (1986).

[17] A. G. Aronov and Yu. V. Sharvin, "Magnetic flux effects in disordered conductors", Reviews of Modern Physics, **59**:755, (1987).

[18] R. Kubo, "A general expression for conductivity tensor", Journal of the Physical Society of Japan, **12**:570, (1957).

[19] D. A. Greenwood, "The Boltzmann equation in the theory of electrical conduction in metals", Proceedings of the Physical Society of London, **71**:585, (1958).

[20] L. Smrčka and P. Středa, "Transport coefficients in strong magnetic fields", Journal of Physics C, **10**:2153, (1977).

[21] J. Kucera and P. Středa, "The relation between transport coefficients and scattering matrices in high magnetic fields", Journal of Physics C, **21**:4357, (1988).

[22] R. Landauer, "Spatial variation of currents and fields due to localized scatterers in metallic conduction", IBM Journal of Research and Development, **1**:223, (1957).

[23] R. Landauer, "Electrical resistance of disordered one-dimensional lattices", Philosophical Magazine, **21**:863, (1970).

[24] M. Büttiker, Y. Imry, R. Landauer, and S. Pinhas, "Generalized many-channel conductance formula with application to small rings", Physical Review B, **31**:6207, (1985).

[25] P. W. Anderson, D. J. Thouless, E. Abrahams, and D. S. Fisher, "New method for scaling theory of localization", Physical Review B, **22**:3519, (1980).

[26] H. L. Engquist and P. W. Anderson, "Definition and measurement of the electrical and thermal resistance", Physical Review B, **24**:1151, (1981).

[27] D. C. Langreth and E. Abrahams, "Derivation of the Landauer conductance formula", Physical Review B, **24**:2978, (1981).

[28] M. Ya. Azbel, "Quantum δ-dimensional Landauer formula", Journal of Physics C, **14**:L225, (1981).

[29] E. N. Economou and C. M. Soukoulis, "Static conductance and scaling theory of localization in one dimension", Physical Review Letters, **46**:618, (1981).

[30] D. S. Fisher and P. A. Lee, "Relation between the conductivity and transmission matrix", Physical Review B, **23**:6851, (1981).

[31] Y. Imry, "Physics of mesoscopic systems", in *Directions in Condensed Matter Physics*, G. Grinstein and G. Mazenko, editors, pages 101–163, World Scientific Press, Singapore, (1986).

[32] S. Washburn, C. P. Umbach, R. B. Laibowitz, and R. A. Webb, "Magnetoresistance of small, quasi-one-dimensional, normal-metal rings and lines", Physical Review B, **30**:4048, (1984).

[33] A. B. Fowler, A. Hartstein, and R. A. Webb, "Conductance in restricted dimensionality accumulation layers", Physical Review Letters, **48**:196, (1982).

[34] A. D. Stone, "Magnetoresistance fluctuations in mesoscopic rings and wires", Physical Review Letters, **54**:2692, (1985).

[35] M. Ya. Azbel, "Resonant tunneling and localization spectroscopy", Solid State Communications, **45**:527, (1983).

[36] B. L. Al'tshuler, A. G. Aronov, D. E. Khmel'nitskii, and A. I. Larkin, "Coherent effects in disordered conductors", in *Quantum Theory of Solids*, I. M. Lifshits, editor, MIR, Moscow, (1982).

[37] P. A. Lee and T. V. Ramakrishnan, "Disordered electronic systems", Reviews of Modern Physics, **57**:287, (1985).

[38] Y. Imry, "Active transmission channels and universal conductance fluctuations", Europhysics Letters, **1**:249, (1986).

[39] B. L. Al'tshuler and B. I. Shklovskii, "Repulsion of energy levels and conductivity of small metal samples", Soviet Physics–JETP, **64**:127, (1986).

[40] P. A. Lee and A. D. Stone, "Universal conductance fluctuations in metals", Physical Review Letters, **55**:1622, (1985).

[41] B. L. Al'tshuler, "Fluctuations in the extrinsic conductivity of disordered conductors", JETP Letters, **41**:648, (1985).

[42] P. A. Lee, A. D. Stone, and H. Fukuyama, "Universal conductance fluctuations in metals: Effects of finite temperature, interactions, and magnetic field", Physical Review B, **35**:1039, (1987).

[43] B. L. Al'tshuler and D. E. Khmel'nitskii, "Fluctuation properties of small conductors", JETP Letters, **42**:359, (1986).

[44] W. J. Skocpol, P. M. Mankiewich, R. E. Howard, L. D. Jackel, T. M. Tennant, and A. D. Stone, "Universal conductance fluctuations in silicon inversion layer nanostuctures", Physical Review Letters, **56**, (1987).

[45] A. D. Benoit, S. Washburn, C. P. Umbach, R. B. Laibowitz, and R. A. Webb, "Asymmetry in the magnetoconductance of metal wires and loops", Physical Review Letters, **57**:1765, (1986).

[46] A. D. Benoit, S. Washburn, C. P. Umbach, and R. A. Webb, "Length-independent voltage fluctuations in small devices", Physical Review Letters, **58**:2343, (1987).

[47] W. J. Skocpol, P. M. Mankiewich, R. E. Howard, L. D. Jackel, T. M. Tennant, and A. D. Stone, "Non-local potential measurements of quantum conductors", Physical Review Letters, **58**:2347, (1987).

[48] H. B. G. Casimir, "On Onsager's principle of microscopic reversibility", Reviews of Modern Physics, **17**:343, (1945).

[49] R. Spal, "A new method for measuring the magnetoconductivity tensor of anisotropic crystals", Journal of Applied Physics, **51**:4221, (1980).

[50] M. Büttiker, "Four-terminal phase-coherent conductance", Physical Review Letters, **57**:1761, (1986).

[51] M. Büttiker, "Symmetry in electrical conduction", IBM Journal of Research and Development, **32**:317, (1988).

[52] M. Büttiker, "Voltage fluctuations in small conductors", Physical Review B, **35**:4123, (1987).

[53] H. U. Baranger, A. D. Stone, and D. P. DiVincenzo, "Resistance correlations in multiprobe microstructures", Physical Review B, **37**:6521, (1988).

[54] C. L. Kane, P. A. Lee, and D. P. DiVincenzo, "Voltage fluctuations in multilead devices", Physical Review B, **38**:2995, (1988).

[55] S. Hershfield and V. Ambegaokar, "Resistance fluctuations in a four probe geometry with infinite leads", Physical Review B, **38**:7909, (1988).

[56] A. D. Stone and A. Szafer, "What is measured when you measure a resistance?–Landauer formula revisited", IBM Journal of Research and Development, **32**:384, (1988).

[57] H. U. Baranger and A. D. Stone, "Electrical linear-response theory in an arbitrary magnetic field: A new Fermi-surface formation", Physical Review B, **40**:8169, (1989).

[58] B. J. van Wees, H. van Houten, C. W. J. Beenakker, J. G. Williamson, L. P. Kouwenhoven, D. van der Marel, and C. T. Foxon, "Quantized conductance of point-contacts in a two-dimensional electron gas", Physical Review Letters, **60**:848, (1988).

[59] D. A. Wharam, T. J. Thornton, R. Newbury, M. Pepper, D. A. Ritchie, and G. A. C. Jones, "One-dimensional transport and the quantization of the ballistic resistance", Journal of Physics C, **21**:L209, (1988).

[60] Yu. V. Sharvin, "A possible method for studying Fermi surface", Soviet Physics–JETP, **21**:655, (1965).

[61] A. G. M. Jansen, A. P. van Gelder, and P. Wyder, "Point-contact spectroscopy in metals", Journal of Physics C, **13**:6073, (1980).

[62] K. von Klitzing, "The quantized Hall effect", Reviews of Modern Physics, **58**:519, (1986).

[63] H. van Houten, C. W. J Beenakker, and B. van Wees, "Quantum point contacts", in *Semiconductors and Semimetals*, M. A. Reed, editor, Academic Press, New York, (1990).

[64] G. Timp, A. M. Chang, P. Mankiewich, R. Behringer, J. E. Cunningham, T. Y. Chang, and R. E. Howard, "Quantum transport in an electron waveguide", Physical Review Letters, **59**:732, (1987).

[65] G. Binnig and H. Rohrer, "Scanning tunneling microscopy", Helvetica Physica Acta, **55**:726, (1982).

[66] G. Binnig and H. Rohrer, "Scanning tunneling microscopy: From birth to adolescence", Reviews of Modern Physics, **59**:615, (1987).

[67] J. Tersoff and D. R. Hamann, "Theory and application for the scanning tunneling microscope", Physical Review Letters, **50**:1398, (1983).

[68] J. Bardeen, "Tunneling from a many-particle point of view", Physical Review Letters, **6**:57, (1961).

[69] J. K. Gimzewski and R. Möller, "Transition from the tunneling regime to point contact studied using scanning tunneling microscopy", Physical Review B, **36**:1284, (1987).

[70] J. K. Gimzewski, R. Möller, D. W. Pohl, and R. R. Schlitter, "Transition from tunneling to point contact investigated by scanning tunneling microscopy and spectroscopy", Surface Science, **181**:267, (1987).

[71] V. H. Hallmark, S. Chiang, J. F. Farbolt, J. D. Swalen, and R. J. Weilson, "Observation of atomic corrugation on Au(111) by scanning tunneling microscopy", Physical Review Letters, **59**:2879, (1987).

[72] J. Wintterlin, J. Wiechers, H. Brune, T. Gritsch, H. Höfer, and R. J. Behm, "Atomic-resolution imaging of close-packed metal surfaces by scanning tunneling microscopy", Physical Review Letters, **62**:59, (1989).

[73] H. W. Fink, "Mono-atomic tips for scanning tunneling microscopy", IBM Journal of Research and Development, **30**:460, (1986).

[74] H. W. Fink, "Point sources for ions and electrons", Physica Scripta, **38**:260, (1988).

[75] R. Landauer, "Electrical transport in open and closed systems", Zeitschrift fur Physics B–Condensed Matter, **68**:217, (1987).

[76] M. C. Payne, "Electrostatic and electrochemical potentials in quantum transport", Journal of Physics: Condensed Matter, **1**:4931, (1990).

[77] R. Landauer, "Conductance determined by transmission: Probes, and quantized constriction resistance", Journal of Physics: Condensed Matter, **1**:8099, (1989).

[78] D. J. Thouless, "Why Landauer's formula for resistance is right?", Physical Review Letters, **47**:972, (1981).

[79] M. Büttiker, *NATO ASI on "Granular Nanoelectronics"* held at Il Ciocco, Italy, (1990).

[80] M. Yosefin and M. Kaveh, "Conduction in curvilinear coordinates: Generalization of the Landauer formula", Physical Review Letters, **64**:2819, (1990).

[81] L. I. Glazman, G. B. Lesovik, D. E. Khmel'nitskii, and R. I. Shekter, "Reflectionless quantum transport and fundamental ballistic-resistance steps in microscopic constrictions", JETP Letters, **48**:238, (1988).

[82] G. Kirczenow, "Theory of the conductance of ballistic quantum channels", Solid State Communications, **68**:715, (1988).

[83] Y. Isawa, "Quantized resistance of metallic narrow channels in ballistic regime", Journal of the Physical Society of Japan, **57**:3457, (1988).

[84] A. Szafer and A. D. Stone, "Theory of quantum conduction through a constriction", Physical Review Letters, **62**:300, (1989).

[85] E. G. Haanapel and D. van der Marel, "Conductance oscillations in two-dimensional Sharvin point contacts", Physical Review B, **39**:5484, (1989).

[86] L. Escapa and N. García, "On the ballistic conductance of small contacts and its resonance structure: Trumpet effect washes out resonant structure", Journal of Physics: Condensed Matter, **1**:2125, (1989).

[87] S. He and S. Das Sarma, "Quantum conduction in narrow constrictions", Physical Review B, **40**:3379, (1989).

[88] E. Tekman and S. Ciraci, "Novel features of quantum conduction in a constriction", Physical Review B, **39**:8772, (1989).

[89] A. Yacoby and Y. Imry, "Quantization of the conductance of ballistic point contacts beyond the adiabatic approximation", Physical Review B, **41**:5341, (1990).

[90] Y. Avishai and Y. B. Band, "Ballistic electronic conductance of a wide orifice", Physical Review B, **41**:3253, (1990).

[91] I. B. Levinson, "Quantum-mechanical conductivity of a ballistic point contact", JETP Letters, **48**:301, (1988).

[92] A. Kawabata, "Theory of ballistic transport through a constriction: Quantization of conductance", Journal of the Physical Society of Japan, **58**:372, (1989).

[93] N. García and L. Escapa, "Elastic oscillatory resistances of small contacts", Applied Physics Letters, **54**:1418, (1989).

[94] E. Tekman and S. Ciraci, "Effects of the constriction geometry on quasi-one-dimensional transport: Adiabatic evolution and resonant tunneling", Physical Review B, **40**:8559, (1989).

[95] G. Kirczenow, "Resonant conduction in ballistic quantum channels", Physical Review B, **39**:10452, (1989).

[96] P. F. Bagwell and T. P. Orlando, "Landauer's conductance formula and its generalization to finite voltages", Physical Review B, **40**:1456, (1989).

[97] D. van der Marel and E. G. Haanapel, "Model calculations of the quantum ballistic transport in two-dimensional constriction-type microstructures", Physical Review B, **39**:7911, (1989).

[98] E. Tekman and S. Ciraci, "Ballistic transport through a quantum point contact: Elastic scattering by impurities", Physical Review B, **42**:in press, (1990).

[99] C. S. Chu and R. S. Sorbello, "Effect of impurities on the quantized conductance of narrow channels", Physical Review B, **40**:5941, (1989).

[100] P. F. Bagwell, "Evanescent modes and scattering in quasi-one-dimensional wires", Physical Review B, **41**:10354, (1990).

[101] J. Mašek, P. Lipavsky, and B. Kramer, "Coherent-potential approximation for the zero-temperature DC conductance of weakly disordered narrow systems", Journal of Physics: Condensed Matter, **1**:6395, (1989).

[102] C. W. J. Beenakker and H. van Houten, "Magnetotransport and nonadditivity of point-contact resistances in series", Physical Review B, **39**:10445, (1989).

[103] E. Castaño and G. Kirczenow, "Theory of conductance of parallel ballistic constrictions", Physical Review B, **41**:5055, (1990).

[104] S. E. Ulloa, E. Castaño, and G. Kirczenow, "Ballistic transport in a novel one-dimensional superlattice", Physical Review B, **41**:12350, (1990).

[105] E. Castaño and G. Kirczenow, "Numerical study of ballistic conduction through a constriction with a barrier", Solid State Communications, **70**:801, (1989).

[106] I. B. Levinson, "Potential distribution in a quantum point contact", Soviet Physics–JETP, **68**:1257, (1989).

[107] G. Kirczenow, "Resonant conductance in ballistic quantum channels", Physical Review B, **39**:10452, (1989).

[108] S. E. Laux, D. J. Frank, and F. Stern, "Quasi-one-dimensional electron states in a split gate GaAs/AlGaAs heterostructure", Surface Science, **196**:101, (1988).

[109] D. A. Wharam, U. Ekenberg, M. Pepper, D. G. Hasko, H. Ahmed, J. E. F. Frost, D. A. Ritchie, D. C. Peacock, and G. A. C. Jones, "Empirical relation between gate voltage and electrostatic potential in the one-dimensional electron gas of a split-gate device", Physical Review B, **39**:6283, (1989).

[110] B. J. van Wees, L. P. Kouwenhoven, E. M. M. Willems, C. J. P. M. Harmans, J. E. Mooij, H. van Houten, C. W. J. Beenakker, J. G. Williamson, and C. T. Foxon, "Quantum ballistic and adiabatic electron transport, studied by quantum point contacts", Physical Review B, **42**:to be published, (1990).

[111] T. Ando, A. B. Fowler, and F. Stern, "Electronic properties of two-dimensional systems", Reviews of Modern Physics, **54**:437, (1982).

[112] G. Timp, *NATO ARW on "Systems of Lower Dimensionality"* held at Marmaris, Turkey, (1990).

[113] J. H. Davies and J. A. Nixon, "Fluctuations in submicrometer semiconducting devices caused by the random positions of dopants", Physical Review B, **39**:3423, (1989).

[114] L. P. Kouwenhoven, B. J. van Wees, C. J. P. Harmans, J. G. Williamson, H. van Houten, C. W. J. Beenakker, C. T. Foxon, and J. J. Harris, "Nonlinear conductance of quantum point contacts", Physical Review B, **39**:8040, (1989).

[115] L. I. Glazman and A. V. Khaetski, "Nonlinear quantum conductance of a point contact", JETP Letters, **48**:592, (1988).

[116] E. Castaño and G. Kirczenow, "Theory of nonlinear transport in narrow ballistic constrictions", Physical Review B, **41**:3874, (1990).

[117] M. Büttiker, "Quantized transmission of a saddle point constriction", Physical Review B, **41**:7906, (1990).

[118] E. Tekman and S. Ciraci, "Resonances in laterally confined systems: Quantum point contacts and electron emitting sharp tips", in *Resonant Tunneling in Semiconductors: Physics and Applications*, L. L. Chang, C. Tejedor, and E. E. Mendez, editors, Plenum Press, New York, (in press).

[119] E. Tekman and S. Ciraci, "Adiabatic evolution and resonant tunneling through a one-dimensional constriction", in *Scanning Tunneling Microscopy and Related Methods*, J. Behm, N. García, and H. Rohrer, editors, pages 157–161, Elsevier, Amsterdam, (1990).

[120] A. Kumar, S. E. Laux, and F. Stern, "Channel sensitivity to gate roughness in a split-gate GaAs-AlGaAs heterostucture", Applied Physics Letters, **54**:1270, (1989).

[121] T. J. Thornton, M. L. Roukes, A. Scherer, and B. P. van der Gaag, "Boundary scattering in quantum wires", Physical Review Letters, **63**:2128, (1989).

[122] J. Faist, P. Guéret, and H. Rothuizen, "Possible observation of impurity effects on conductance quantization", Physical Review B, **42**:3217, (1990).

[123] I. Kander, Y. Imry, and U. Sivan, "Effects of channel opening and disorder on the conductance of narrow wires", Physical Review B, **41**:12941, (1990).

[124] K. S. Ralls and R. A. Buhrman, "Defect interactions and noise in metallic nanoconstrictions", Physical Review Letters, **60**:2434, (1988).

[125] D. A. Wharam, M. Pepper, H. Ahmed, J. E. F. Frost, D. G. Hasko, D. C. Peacock, D. A. Ritchie, and G. A. C. Jones, "Addition of the one-dimensional quantized ballistic resistance", Journal of Physics C, **21**:L887, (1988).

[126] C. G. Smith, M. Pepper, H. Ahmed, J. E. F. Frost, D. G. Hasko, D. C. Peacock, D. A. Ritchie, and G. A. C. Jones, "Transition from one- to zero-dimensional ballistic transport", Journal of Physics C, **21**:L893, (1988).

[127] E. Tekman, *"Atomic Theory of the Scanning Tunneling Microscope"*, Master's thesis, Bilkent University, Ankara, Turkey, (1988).

[128] E. Tekman and S. Ciraci, "Tip induced localized states in scanning tunneling microscopy", Physica Scripta, **38**:486, (1988).

[129] E. Tekman and S. Ciraci, "Atomic theory of scanning tunneling microscopy", Physical Review B, **40**:10286, (1989).

[130] U. Dürig, J. K. Gimzewski, and D. Pohl, "Experimental observation of forces acting during scanning tunneling microscopy", Physical Review Letters, **57**:2403, (1986).

[131] N. D. Lang, "Theory of single-atom imaging in the scanning tunneling microscope", Physical Review Letters, **56**:1164, (1986).

[132] N. D. Lang, "Resistance of a one-atom contact in the scanning tunneling microscope", Physical Review B, **36**:8173, (1987).

[133] N. D. Lang, "Apparent barrier height in scanning tunneling microscope", Physical Review B, **37**:10395, (1988).

[134] S. Ciraci, A. Baratoff, and I. P. Batra, "Tip-sample interaction effects in scanning-tunneling and atomic-force microscopy", Physical Review B, **41**:2763, (1990).

[135] S. Ciraci, A. Baratoff, and I. P. Batra, "Site-dependent electronic effects, and deformations in scanning tunneling microscopy of flat metal surfaces", Physical Review B, **42**:in press, (1990).

[136] S. Ciraci, E. Tekman, A. Baratoff, and I. P. Batra, (preprint).

[137] J. Ferrante and J. R. Smith, "A theory of adhesion at a bimetallic interface: Overlap effects", Surface Science, **38**:77, (1973).

[138] N. García, *STM '87 Conference* held at Oxnard, California, (1987).

[139] J. Ferrer, A. Martin-Rodero, and F. Flores, "Contact resistance in the scanning tunneling microscope at very small distances", Physical Review B, **38**:10113, (1988).

[140] S. Ciraci and E. Tekman, "Theory of transition from tunneling regime to point contact in scanning tunneling microscopy", Physical Review B, **40**:11969, (1989).

[141] U. Dürig, O. Züger, and D. Pohl, "Observation of metallic adhesion using the scanning tunneling microscope", Physical Review Letters, **65**:349, (1990).

[142] J. R. Smith, "Self-consistent many-electron theory of electron work functions and surface potential characteristics for selected metals", Physical Review, **181**:522, (1969).

[143] W. A. Harrison, *Electronic Structure of and the Properties of Solids*, Freeman, San Fransisco, (1980).

[144] N. J. Zheng and I. S. T. Tsong, "Resonant-tunneling theory of imaging close-packed metal surfaces by scanning tunneling microscopy", Physical Review B, **41**:2671, (1990).

[145] J. C. Chen, "Origin of atomic resolution on metal surface in scanning tunneling microscopy", Physical Review Letters, **65**:448, (1990).

[146] E. Tekman and S. Ciraci, "Theory of anomalous corrugation of the Al(111) surface obtained from scanning tunneling microscopy", Physical Review B, **42**:1860, (1990).

[147] Vu Thien Binh and J. Marieu, "Characterization of microtips for scanning tunneling microscopy", Surface Science, **202**:L539, (1989).

[148] N. García, J. J. Sáenz, and H. De Raedt, "Electron emission from small sources", Journal of Physics: Condensed Matter, **1**:9931, (1989).

[149] N. D. Lang, A. Yacoby, and Y. Imry, "Theory of a single-atom point source for electrons", Physical Review Letters, **63**:1499, (1989).

[150] H. De Raedt, N. García, and J. J. Sáenz, "Focused electron emission from planar quantum point contacts", Physical Review Letters, **63**:2260, (1989).

[151] J. Summhammer and J. Schmiedmayer, "Wave aspects of electron and ion emission from point sources", Physica Scripta, **40**:in press, (1990).

[152] E. Tekman, S. Ciraci, and A. Baratoff, "Theoretical study of collimated field emission of electrons from a point source", Physical Review B, **42**:in press, (1990).

[153] L. Orosz and E. Balázs, "Calculation of a quasi-self-consistent barrier for metal-vacuum-metal junctions", Surface Science, **177**:444, (1986).

[154] P. A. Serena, L. Escapa, J. J. Sáenz, N. García, and H. Rohrer, "Coherent electron emission from point sources", Journal of Microscopy, **152**:43, (1988).

[155] M. L. Roukes, A. Scherer, S. J. Allen, Jr., H. G. Craighead, R. M. Ruther, E. D. Beebe, and J. B. Harbison, "Quenching of the Hall effect in a one-dimensional wire", Physical Review Letters, **59**:3011, (1987).

[156] H. U. Baranger and A. D. Stone, "Quenching of the Hall resistance in ballistic microstructures: A collimation effect", Physical Review Letters, **63**:414, (1989).

[157] C. J. B. Ford, S. Washburn, M. Büttiker, C. M. Knoedler, and J. M. Hong, "Influence of geometry on the Hall effect in ballistic wires", Physical Review Letters, **62**:2724, (1989).

[158] A. M. Chang, T. Y. Chang, and H. U. Baranger, "Quenching of the Hall resistance in a novel geometry", Physical Review Letters, **63**:996, (1989).

[159] L. W. Molenkamp, A. A. M. Staring, C. W. J. Beenakker, R. Eppenga, C. E. Timmering, J. G. Williamson, C. J. P. M. Harmans, and C. T. Foxon, "Electron-beam collimation with a quantum point contact", Physical Review B, **41**:1274, (1990).

[160] M. L. Roukes, A. Scherer, and B. P. Van der Gaag, "Are transport anomalies in "Electron waveguides" classical?", Physical Review Letters, **64**:1154, (1990).

[161] C. W. J. Beenakker and H. van Houten, "Billiard model of a ballistic multiprobe conductor", Physical Review Letters, **63**:1857, (1989).

[162] L. P. Kadanoff and G. Baym, *Quantum Statistical Mechanics*, W. A. Benjamin, New York, (1962).

[163] L. V. Keldysh, "Diagram technique for nonequilibrium processes", Soviet Physics–JETP, **20**:1018, (1965).

[164] R. A. Craig, "Perturbation expansion for real-time Green's functions", Journal of Mathematical Physics, **9**:605, (1968).

[165] J. Rammer and H. Smith, "Quantum field-theoretical methods in transport theory of metals", Reviews of Modern Physics, **58**:323, (1986).

[166] G. D. Mahan, "Quantum transport equation for electric and magnetic fields", Physics Reports, **145**:251, (1987).

[167] S. Datta, "Steady-state quantum kinetic equations", Physical Review B, **40**:5830, (1989).

[168] M. Büttiker, "Voltage fluctuations in small conductors", Physical Review B, **35**:4123, (1987).

[169] M. Büttiker, "Role of quantum coherence in series resistors", Physical Review B, **33**:3020, (1986).

[170] M. Büttiker, "Coherent and sequential tunneling in series barriers", IBM Journal of Research and Development, **32**:63, (1988).

[171] J. R. Kirtley, S. Washburn, and M. J. Brady, "Scanning tunneling measurements of potential steps at grain boundaries in the presence of current flow", Physical Review B, **32**:414, (1988).

[172] C. S. Chu and R. S. Sorbello, "Scanning tunneling microscope as a probe of the local transport field in mesoscopic systems", Physical Review B, **40**:5950, (1989).

[173] M. Büttiker, "Chemical potential oscillations near a barrier in the presence of transport", Physical Review B, **40**:3409, (1989).

[174] J. L. D'Amato and H. M. Pastawski, "Conductance of a disordered chain including inelastic scattering events", Physical Review B, **41**:7411, (1990).

[175] P. L. Pernas, A. Martin-Rodero, and F. Flores, "Electrochemical-potential oscillations across a constriction", Physical Review B, **41**:8553, (1990).

[176] S. Feng, "Quantum transport in the presence of phase breaking scattering: Generalized Landauer formula", Physics Letters A, **143**:400, (1990).

Vita

A. Erkan Tekman was born in Ankara, on 29 April 1965. He attended to the Electrical and Electronics Engineering Department, Middle East Technical University in September 1982 and graduated with high honors in July 1986. In September 1986 he joined to Department of Physics, Bilkent University as a Research Assistant to carry out graduate study with Prof. Salim Çıracı. In Master's study he worked on "Atomic Theory of the Scanning Tunneling Microscope", and got his degree in January 1988. His later interest was on the ballistic transport through quantum point contacts and the effect of lateral confinement on tunneling, especially in scanning tunneling microscopy. His research activities also include atomic force microscopy, semiconductor heterostructures and Schottky barriers. He was appointed as an Instructor at Department of Physics, Bilkent University in October 1988. He spent three months in Zurich Research Laboratory, IBM Research Division in 1989 as a visiting scientist, working mainly on the electron emission properties of point sources and quantum point contacts. He published articles in Solid State Communications, Physica Scripta, Journal of Physics: Condensed Matter and Physical Review B; and delivered lectures and talks in NATO Advanced Study Institutes and Research Workshops. He is currently working on theory of quantum transport and scanning tunneling microscopy and is an Assistant Professor at Department of Physics, Bilkent University.